

**TEACHER LEADER
BEHAVIOR AND ITS RELATION
TO TEACHER EFFECTIVENESS**

For Reference

NOT TO BE TAKEN FROM THIS ROOM

By

Arthur Groat McBeath

1959

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS





Digitized by the Internet Archive
in 2019 with funding from
University of Alberta Libraries

https://archive.org/details/McBeath1959_0

Thesis
1959(F)
#29

TEACHER LEADER BEHAVIOR AND ITS RELATION
TO TEACHER EFFECTIVENESS

A Thesis

Submitted to

the Division of Educational Administration

the Faculty of Graduate Studies

The University of Alberta

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

by

Arthur Groat McBeath

August 1959

UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Teacher Leader Behavior and Its Relation to Teacher Effectiveness" submitted by Arthur Groat McBeath in partial fulfillment of the requirements for the degree of Master of Education.

ABSTRACT

This thesis studied the relationship between teacher leader behavior in the classroom and ratings of teacher effectiveness. Leader behavior dimensions of Consideration and Initiating Structure were considered the crucial variables within a theoretical framework of variables for studying teacher effectiveness.

In 12 schools in Alberta 59 teachers in grades 8 to 12 were the centre of the study. They were rated on teaching effectiveness and their leader behavior was described by themselves, by 571 students, 50 other teachers on the same staff, 12 principals, and 12 superintendents. The Teacher Leader Behavior Description Questionnaire was used to describe the teacher's leader behavior and a global scale was used to give a single rating of his teaching effectiveness.

The study supported the following conclusions:

1. The most effective teachers were those who were described as being above the average on Consideration and Initiating Structure behavior. The least effective teachers were those who were described as being below the average on both dimensions of leader behavior.

2. Students related effective teaching to both

Consideration and Initiating Structure behavior. There was a tendency for Consideration behavior to be more closely related to ratings of effectiveness.

3. Administrators related effective teaching to both Consideration and Initiating Structure behavior, but neither dimension was more closely related to ratings of effectiveness.

Summing up, one of the factors related to teacher effectiveness in the classroom was the extent to which the teacher exhibited human relations behavior and behavior which attempted to get things done or accomplish goals. This appeared true in teaching grades eight to twelve, in both academic and non-academic subjects, and in a wide variety of non-urban schools.

ACKNOWLEDGMENTS

The writer wishes to express his thanks to his advisor, Dr. J. H. M. Andrews, and to the other members of his thesis committee, Dr. R. S. MacArthur and Dr. A. W. Reeves, for their valuable guidance and constructive criticism.

Special thanks are extended to the students, teachers, principals, and superintendents who took part in this research. Without their outstanding co-operation the study obviously would have been impossible.

Appreciation is also expressed to the W. K. Kellogg Foundation and the Alberta Advisory Committee on Educational Research for their financial assistance.

TABLE OF CONTENTS

	PAGE
LIST OF TABLES	x
LIST OF FIGURES	xiii
CHAPTER	
I. THE PROBLEM	1
THE GENERAL PROBLEM	1
NEED FOR STUDY OF TEACHER EFFECTIVENESS	2
LIMITATION OF THE STUDY	4
SPECIFIC PROBLEMS	5
II. RELATED LITERATURE	7
APPROACHES TO THE STUDY OF TEACHER	
EFFECTIVENESS	7
Teacher Effectiveness Studied through the	
Results Produced in Pupils	7
Teacher Effectiveness Studied through the	
Traits of Teachers	9
Teacher Effectiveness Studied through the	
Performance of Teachers	12
Teacher Effectiveness Studied through the	
Behavior of Teachers	13
EVALUATING TEACHER EFFECTIVENESS	16
PERCEPTION OF BEHAVIOR AND JUDGMENT	
OF EFFECTIVENESS	21

CHAPTER	PAGE
SUMMARY	23
III. THEORY AND METHODOLOGY	26
FRAMEWORK TO STUDY TEACHER EFFECTIVENESS . .	26
TEACHER BEHAVIOR	28
JUDGES OF TEACHER LEADER BEHAVIOR AND	
TEACHER EFFECTIVENESS	31
HYPOTHESES	33
PROCEDURE FOR TESTING HYPOTHESES	34
Instruments	35
Global rating scale of teacher	
effectiveness	35
The Ohio Leader Behavior Description	
(Questionnaire) (LBDQ)	37
The Teacher Leader Behavior Description	
Questionnaire (TLBDQ)	39
The Pilot Study	41
Plan for Collecting the Data	42
SUMMARY	43
IV. COLLECTION OF THE DATA	47
METHOD OF COLLECTING THE DATA	47
The Teachers Who Were the Centre of	
the Study	47
The Other Teachers on the Same Staff . . .	48

CHAPTER	PAGE
The Students	48
The Principals	50
The Superintendents	50
THE SCHOOLS AND TEACHERS INVOLVED	51
SUMMARY	52
V. RESULTS, ANALYSIS AND DISCUSSION OF	
EFFECTIVENESS RATINGS	54
RESULTS AND ANALYSIS	54
Group Averages and Deviations	54
Consensus.	59
Variations by Schools from the Average	
for the Scale.	63
DISCUSSION	64
VI. RESULTS, ANALYSIS AND DISCUSSION OF	
DESCRIPTIONS OF TEACHER LEADER BEHAVIOR. . .	68
RESULTS AND ANALYSIS	68
Group Averages and Deviations	68
Independence of the Dimensions	72
Consensus.	74
DISCUSSION	79
VII. THE RELATIONSHIP BETWEEN EFFECTIVENESS RATINGS	
AND DESCRIPTIONS OF LEADER BEHAVIOR.	86
RESULTS AND ANALYSIS	86

CHAPTER	PAGE
High and Low Effectiveness Groups	86
Rank-order Correlations	96
DISCUSSION.	104
VIII. CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS .	109
CONCLUSIONS	110
IMPLICATIONS	113
RECOMMENDATIONS	116
BIBLIOGRAPHY	118
APPENDIX A	121
APPENDIX B	127

LIST OF TABLES

TABLE		PAGE
I.	Medians, Quartile Points and Quartile Deviations of the Effectiveness Ratings for Fifty-nine Teachers	55
II.	Frequency of Extreme Scores Used in Rating Fifty-nine Teachers	58
III.	Rank-order Correlations between Raters	60
IV.	Consensus among Judges about Effectiveness Ratings	62
V.	Medians, Quartile Points and Quartile Devi- ations of the Behavior Descriptions for Fifty-nine Teachers	69
VI.	Correlations between Initiating Structure and Consideration Scores for Each Set of Judges .	72
VII.	Correlations between Judges on the Initiating Structure Scores	75
VIII.	Correlations between Judges on the Consideration Scores.	75
IX.	Consensus among Judges about Behavior Descriptions	78
X.	Number of Teachers in High and Low Effective- ness Groups Scoring above the Median on Both	

TABLE

PAGE

	Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions.	88
XI.	Number of Teachers in High and Low Effective- ness Groups Scoring above the Median on Both Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions as Described by Students	91
XII.	Number of Teachers in High and Low Effective- ness Groups Scoring above the Median on Both Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions as Described by Other Teachers.	92
XIII.	Number of Teachers in High and Low Effective- ness Groups Scoring above the Median on Both Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions as Described by Principals.	93
XIV.	Number of Teachers in High and Low Effective- ness Groups Scoring above the Median on Both Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions as Described by Superintendents	94

TABLE	PAGE
XV. Number of Teachers in High and Low Effectiveness Groups Scoring above the Median on Both Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions as Described by Teachers Themselves	95
XVI. Number of Teachers in High and Low Effectiveness Groups Scoring above the Median on Both Leader Behavior Dimensions, and Scoring below the Median on Both Behavior Dimensions as Described by One Half-group of Students.	97
XVII. Correlations between Effectiveness Ratings and Behavior Descriptions by the Same Judge .	98
XVIII. Correlations between Effectiveness Ratings and Behavior Descriptions by Separate Judges.	101

LIST OF FIGURES

FIGURE	PAGE
1. Framework for Research in Teacher Effectiveness. .	29

CHAPTER I

THE PROBLEM

The institution of education is based on the teachers in the classrooms. Improvement of the effectiveness of the teacher is of concern to teachers, administrators, students, parents, trustees and the general public. Studies of teaching have analyzed many aspects of the work of the teacher with the hope that some insight into his work will provide a background for improvement in effectiveness. This thesis is a further study of teacher effectiveness.

I. THE GENERAL PROBLEM

This research was concerned with an attempt at describing specific teacher behavior and the relationship of this behavior to teacher effectiveness. This attempt originated from empirical evidence in the field of leadership which was considered applicable to the field of teaching. Evidence gathered through the Ohio State University Leadership Studies initiated in 1945 has provided a means of describing leader behavior. The teacher in the classroom is in a leadership situation where he is interested in goal achievement and human relations. Thus, the problem studied was the relationship between descriptions of teacher leader behavior and judgments of teacher effectiveness.

II. NEED FOR STUDY OF TEACHER EFFECTIVENESS

When one hears the saying, "A school is no better than its teachers.", the importance of the teacher is well illustrated. The improvement of our educational system is dependent on the improvement of our training of teachers both in the teacher training institutions and in the field. This improvement of teachers is dependent upon better selection of candidates to eliminate those who will be ineffective. These attempts at improvement are based on the assumption that we are able to define an effective teacher.

That we are not yet able to define the effective teacher is supported by the following statements. The Encyclopedia of Educational Research states,

. . . it is apparent that the identification and definition of teaching competencies is as yet by no means satisfactory. . . . we do not yet have an adequate definition of teaching efficiency. . . .
(4, p. 1453)

Kinney says,

Clearly, then, whether we are administrators, supervisors, teacher educators, placement officers, parents, teachers, students, or merely one among the various categories of critics of the public schools, each of us has his definition of teaching competency--fuzzy or clear, implicit or explicit, narrow or comprehensive, it serves as the basis for significant decisions. Unless we arrive at a common understanding, the result must continue as confusion, controversy, and inefficiency. . . . (3, p. 1)

Ackerman states, ". . . there is hardly any conclusive evidence as to the nature and means of identifying teacher competence." (1, p. 273)

We have assumed that there is a need to improve teacher effectiveness. An attempt at improvement first requires an evaluation of how effective teachers are at present. But without a definition or an understanding of teacher effectiveness, an evaluation cannot take place. Thus, improvement of teaching is restricted.

Without an understanding of what makes a teacher effective, we have difficulty in preparing the most suitable teacher training program. Ryans believes that we are certain that we want as many good teachers as possible. To obtain these teachers we have created systems for selecting teachers, have devised programs of training, and set up systems of supervision to improve teaching. These all have been based on common sense and practical wisdom with little scientific foundation, but with high hopes and deep conviction that we are doing the best (5, p. 238). The need for a study of teacher effectiveness is apparent if we desire to improve our training of teachers.

Our supervision of teachers implies that the supervisor desires to improve instruction, but without some understanding of what makes an effective teacher much of our supervisory effort must be wasted. Burton and Brueckner

state,

. . . All in all, teaching is a very complex activity, and the haphazard, unscientific, and superficial study of teaching that characterizes much of our supervision today should not be tolerated. . . . Just as we have developed improved methods of studying pupils and their habits of work, so we must develop improved methods of studying and assisting teachers.
(2, p. 359)

III. LIMITATION OF THE STUDY

The attempt to solve the problem of teacher effectiveness has been characterized by two lines of attack. One has been the traits approach, assuming that effectiveness is dependent on the individual teachers. The other, which has received less emphasis, has been the situational approach, assuming effectiveness is dependent on the situation in which the teacher is placed.

This study assumed that both the behavior of the individual teacher and the situation in which he is placed are related to effectiveness. The study was limited, though, by being concerned only with the behavior of the teachers no matter how variable were the situations in which they taught.

The study was not concerned with character traits of teachers since research on evaluation of teacher effectiveness has shown this approach to be fruitless. The study was concerned with teacher leader behavior in the classroom.

No attempt was made to relate leader behavior to the type of school system, subjects or grades taught, experience and training of the teachers, type of community or any of the numerous other, situational variables.

Teacher effectiveness was limited to the actual teaching of students in the classroom and was defined as the opinions of five sets of raters. These were students, principals, superintendents, other teachers on the same staff, and the teachers themselves.

IV. SPECIFIC PROBLEMS

1. The major problem was to study the relationship between descriptions of teacher leader behavior and ratings of effectiveness when the descriptions of leader behavior and ratings of effectiveness were provided by students, other teachers, principals, superintendents and the teachers themselves.

2. A secondary problem was concerned with the various viewpoints that the five sets of raters would provide when rating the teacher's effectiveness. Thus a second problem was to study the relationships between the effectiveness ratings provided by these raters.

3. A parallel problem to the second one was to study the relationships between the descriptions of the teacher's leader behavior provided by these five sets of raters.

REFERENCES

- (1) Ackerman, W. I. "Teacher Competence and Pupil Change," Harvard Educational Review, 24:273-289, 1954.
- (2) Burton, W. H., and L. J. Brueckner. Supervision a Social Process. Third Edition. New York: Appleton-Century-Crofts Inc., 1955.
- (3) Kinney, L. B. "What Is a Good Teacher?" School of Education Bulletin, Utah State Agricultural College, Utah, 1954.
- (4) Monroe, W. S. (ed.). Encyclopedia of Educational Research. New York: The MacMillan Co., 1952.
- (5) Ryans, D. G. (chairman), et al. "Report of the Committee on the Criteria of Teacher Effectiveness," Review of Educational Research, 22:238-263, June, 1952.

CHAPTER II

RELATED LITERATURE

This chapter reviews some of the literature related to previous approaches to the solution of the problem of teacher effectiveness, some of the literature related to evaluation of teacher effectiveness, and literature concerning the processes of perceiving and judging the behavior and effectiveness of individuals.

I. APPROACHES TO THE STUDY OF TEACHER EFFECTIVENESS

Four approaches are reported in this section. In the first, teacher effectiveness was studied through the results produced in the pupils. In the second approach, teacher effectiveness was studied through the traits which are characteristic of good and poor teachers. The third approach was to study teacher performance and the final approach reported was the study of teacher behavior.

Teacher Effectiveness Studied through the Results Produced in Pupils

The extreme example of this type of approach was concerned with changes in the pupil as a member of the community. What changes in the members of the community

(the former pupils) were brought about by the teacher? What kind of teachers or teaching brought about this desirable change? In favour of this approach are some sociologists such as Payne who believed we must use devices measuring total behavior changes in family and community to evaluate effectiveness of education (12, p. 670). He states, ". . . results of education should be measured in the community rather than in the school." (12, p. 671) This approach requires agreement firstly, on a definition of a desirable change in the pupil, and secondly, isolation of the effects of the teacher on the pupil as distinct from the effects of other aspects of the community. Both of the above requirements appear very difficult to obtain.

A great deal has been reported using pupil growth and achievement in school as the criterion for effective and ineffective teaching. There has been much disagreement over this type of approach. It has been difficult to isolate the effect of a single teacher and also it has been difficult to control the many pupil variables. Many of the criteria of successful and non-successful pupils have been in the form of standardized tests and although these may indicate some aspects of desirable pupil change, they do not indicate all aspects. Biber believes that it may be possible through intensive psychological studies, using trained personnel

and by controlling such factors as family and socio-economic background, to study pupil growth as a measure of the cumulative effect of education, but not as an index of an individual teacher's effectiveness (3, p. 215). Reavis and Cooper summarize the research using this approach:

. . . one of the major efforts of the past twenty years in teacher evaluation has been the attempt to isolate the influence of teaching ability as a single variable affecting pupil achievement. Age, intelligence, initial achievement and socio-economic status of pupils have been held constant in various significant studies of pupil achievement. Several types of accomplishment quotients and pupil progress indices have resulted from these studies.

The disappointing feature of attempts to relate pupil results to teaching success has been that pupil results have not correlated highly with other measures of teacher success. In particular, the correlations between pupil results and administrator or supervisor ratings have been low. (13, pp. 100, 101)

Ackerman, in reviewing, concludes,

. . . more than thirty studies were reviewed. In each of these studies an attempt was made to evaluate teacher competence and effectiveness by the use of the criterion of pupil change. In general the results are contradictory and inconsistent. (1, p. 284)

Teacher Effectiveness Studied through the Traits of Teachers

The assumption behind this approach has been that if one can determine those traits which are characteristic of good and poor teachers, one would be able to predict which

teachers would be most effective in the classroom, school and community.

A detailed traits approach was used in the Commonwealth Teacher Training Study and reported by Charters and Waples in 1929. About this study, Beecher notes that a master list of twenty-five teacher traits was developed which seemed to indicate nothing unique about teachers (2, p. 13).

Gray, when surveying the present practices in Canada, Great Britain and the United States, found four qualities of teachers most important. They were attitude toward work, personality, spirit in the classroom, and stimulation of initiative (5, p. 171).

Teacher intelligence has been a well studied trait and the lack of conclusive evidence, often a negative correlation, has been disquieting to teacher educators. Beecher's summary describes a comprehensive study under the auspices of the University of Wisconsin and The State Department of Public Instruction in Wisconsin which was carried out from 1934 to 1944. One phase of the study indicated that intelligence of the teacher seemed to be the highest conditioning factor in teacher success, but the conclusion was qualified by stating that the results should be verified (2, p. 27). This was done by Rolfe using the

same procedure and it was found that intelligence showed a small negative correlation with teacher success (2, p. 28).

From her research Morris reported that,

Examination of both temperamental and intellectual traits show, further, that there is probably no unique blend of traits so characteristic of the successful teacher, as markedly different from the successful librarian or other professional representative, that this blend or constellation of traits may in itself be used as the basis for selecting teachers. But there seem to be differences in the subtle 'balance of power' that reflect the varying degrees of certain traits as they are combined in different individuals.
(10, p.2)

If one were able to compile a list of desirable traits of teachers, another problem would arise in trying to determine the relative weights of each trait which would make up the cumulative measure of teacher effectiveness.

Gage and Orleans believe that teacher effectiveness is not to be considered as consisting of any single fixed pattern of traits (4, pp. 294-298). In any given pattern that produces desired effects on pupil behavior, the elements may vary and may be compensatory. Thus, with modifications a different pattern may still produce effective results. Common elements may not be present in effective teachers. This led Orleans, et al. to conclude that looking for common elements in competent teachers was unrealistic. They suggested that the problem was one of being able to predict what kind of teachers would be effective

with what kinds of pupils for promoting change toward what kinds of goals and in what kinds of situations (11, p. 646).

A similar belief was expressed by Ryans in reporting that,

. . . The group also agreed that to define the effective teacher as a unitary concept was impossible and that we had better also forget the concept of an "ideal" teacher. Positively, we should recognize that there may be effective teachers who may differ markedly among themselves. Perhaps the problem should be studied narrowly: "What is effective teaching in French or beginning arithmetic?"
 . . . (16, p. 240)

Thayer expressed doubt about studying a teacher by breaking his general effectiveness into parts when he says, ". . . and the whole of a teacher's effectiveness may be greater or less than the sum of the values assigned to his analyzed parts." (21, p. 370)

Teacher Effectiveness Studied through the Performance of Teachers

The assumption with this approach has been that knowledge of specific performances of teachers which are considered effective will allow accurate prediction with regard to pupil growth. Superiors have rated certain performances of teachers in the classroom as being effective or ineffective. Pupils have indicated what performances they "like" and "dislike."

One fault of this approach has been the combining of the description of the performance with the judgment of that performance as being effective or ineffective. Research that telescopes these two steps never gives a complete picture of what the teacher does. Rather than describing performance and then making a value judgment about the teacher's effectiveness, persons have tended to describe only those actions which they like or dislike.

Teacher Effectiveness Studied through the Behavior of Teachers

Specific reference to "behavior" was found in much of the more recent literature on teacher effectiveness. Behavior as a psychological term refers to the individual's overt or covert reaction to stimuli and there is no connotation of being "good" or "bad."

Mitzel contends that the crucial area for the study of teacher effectiveness, ". . . is not the selection and measurement of appropriate criteria, but the identification, isolation and measurement of the classroom behaviors of pupils and teachers." (9, p. 1) Much study has been concerned with the relationship of individual teacher factors, such as character traits, education and experience, to that of pupil growth while ignoring the teacher's behavior.

Thus, inconsistencies have arisen. Levin gave an example from a study of teacher effectiveness by Davis where teachers' behavior was overlooked and the results were perplexing. In this study the achievement of two groups of secondary school pupils was compared. The pupils in one group had teachers with neither a major nor minor in the subjects they were teaching. In the second group the pupils had teachers with a major or minor in the subjects they were teaching. The former group of pupils scored higher on achievement tests, and one of the reasons could be that the study by-passed the actual behavior of the teachers in the classroom (7, p. 103).

Ackerman says:

Learning is a change in behavior. A teacher is effective when he does things or behaves in ways that engender the learning of skills, understandings, work habits, desirable attitudes and adequate personal adjustment on the part of the pupils or students. Looked at from this point of view neither pupil change nor teacher effectiveness is a unitary concept. Each admits of a number of possibilities.
 . . . (1, p. 284)

.

A trait, characteristic, or quality is a meaningless concept unless it can be anchorable to some sort of denotable behavior. . . . (1, p. 286)

.

The observation of classroom behavior becomes, therefore, a crucial step in the entire process. It is, perhaps, the only valid appraisal of teacher factors that can be made . . . behaviors are the crucial point in the conditioning of pupil change.
 (1, p. 286)

Gage and Orleans (4), Orleans et al. (11), Melby (8) and Hughes (6) all drew attention to the need for the study of teacher behavior patterns in research in the field of teacher effectiveness.

Beecher (2, pp. 64-78) developed a scale for rating effective teacher behavior. The components of the scale were characteristics based on what pupils say they like in teachers. The categories used were indications of fairness, cheerfulness, sympathetic understanding, control, ability to get pupil response, knowledge and skill. The criterion judges were considered experts by their colleagues. The scores on the behavior scale showed a correlation of .88 with ratings of the same teachers by the criterion judges. Beecher believed the correlation indicates a high degree of validity of the instrument. Reliability also appeared high when tested by repeated ratings and also by other raters. The author concluded by stating, ". . . the findings justify the conclusion that valuable appraisal evidence can be obtained on the basis of classroom behavior to which research indicates pupils react most frequently." (2, p.77) There was some indication that the scale recorded teacher characteristics rather than overt behavior. The choice of the areas according to pupil "likes" and "dislikes" restricts effective behavior to impressions of pupils. There was no

indication that the six behavior areas or dimensions were independent. Some of the dimensions may be parts of other dimensions in the set or may be parts of dimensions not included in the set of six.

The evidence above supported the contention that study of teacher behavior might be a more worthwhile approach to the study of teacher effectiveness.

II. EVALUATING TEACHER EFFECTIVENESS

Since an integral part of this study involved the rating of a teacher's effectiveness, some of the literature on this subject is reported.

Ratings provide the most common type of quantitative appraisal of teacher effectiveness. They may be direct comparisons which involve ranking or pairing or they may involve scales. The rating may be based on teacher traits, teacher performances, pupil results, or a combination of these.

In 1921, Rugg (15, p. 425) described some very extensive research concerning the rating of human character. He believed that rating on a point scale was possible if the final rating was an average of at least three independent ratings, if the scales were comparable and equivalent, and if the three raters were competent by being thoroughly acquainted with the individual being rated. This, he believed,

was unattainable in public schools. In work done with army officers during World War 1, he reports, ". . . it was very improbable that an officer was located within even his proper 'fifth' of the entire scale by an 'official' rating. The ratings were practically valueless." (15, p. 436) Later Rugg worked on a more elaborate experiment using a "man-to-man" comparison. Here the rating was made by comparing the person to be rated against a standard determined by a select prior ranking. "The chances can not be more than 4 to 1 that any rating will be within 14 points [out of 80] of the person's true rating." (15, p. 488)

The persons most thoroughly acquainted with the teacher's work are the students. Beecher reported Bryan's study of "Pupil Rating of Secondary School Teachers" which was an early attempt to include the pupils as the raters. He reported the correlation between students' and administrators' ratings of the same teacher was .68 (2, p. 18). Beecher summarized some of the studies, which used the pupil as rater, by noting, ". . . it appears that both junior and senior high school pupils can point out specific strong and weak spots in teaching to a degree that makes pupil ratings worthwhile." (2, p. 18)

The relationship between general ratings of a teacher's effectiveness and pupil results was studied by

McCall. Shane and Yauch reported that McCall's study showed no statistically significant evidence of relationships between pupil progress and ratings by principals, ratings by peers, and self-ratings (18, p. 157).

Criticisms of ratings of teacher effectiveness have been frequent. One area of such criticism has been the standard against which the teacher would be judged. The most common standard has been a fuzzy picture of the "ideal" teacher. Biber believes, however, that it is futile to measure a teacher against a fixed point of the "ideal" teacher. She considers the effective teacher in terms of a possible range within which satisfactory teaching practices may fall. Outside both extremities of this range are practices which are considered unsatisfactory (3, p. 216). As a standard the "ideal" teacher is limited due to the great amount of variability in the definitions of "ideal."

Thayer criticized ratings because they are too superficial. He states that a rating,

. . . serves to give one the first over-view of the teaching performance, a general and superficial job analysis, rather than a vital and profound analysis of the teaching act. . . . (2, p. 376)

Rugg believed that we rate individuals in terms of general mental attitude toward them. This is called the "halo" effect. This mental attitude toward the individual dominates judgment of particular qualities (15, p. 37).

This belief points out the disadvantages of itemized rating scales. Rugg gave an illustration of this theory when he reported on a captain who was rated the lowest of twenty by thirteen officers on all categories of the scale such as physical, intellectual, and leadership qualities. Yet, on an objective intelligence test, the captain scored the highest of the twenty people rated and he had been a Rhode's scholar. The opinions of the rating officers of the captain's general personal quality influenced their opinion on the specific categories (15, pp. 37, 38). A similar criticism has been voiced about students' ratings of teachers. Their sometimes shallow "likes" and "dislikes" influence their rating of the teacher on specific items.

Another area of criticism was the lack of objectivity of the ratings. A common belief is that teaching is an art and many of the results are intangible. Therefore, to rate teaching objectively is difficult, and perhaps impossible. Young notes that, "Any rating scale is largely subjective even when developed to the greatest extent and operated with extreme care." (22, p. 76)

Shane and Yauch state:

. . . It became apparent in a few years that "improved" rating scales were simply a refinement of an original error. If the basis for judgment was to be human subjectivity, it made little difference whether or not the instrument for the collection of evidence was objectively constructed:

the value-judgments remained subjective. After the evidence was in, it still had to be evaluated by fallible humans. (18, p. 141)

Some of the disadvantages of subjective ratings are that they tend to be unreliable, they do not allow the rater to give a detailed explanation for the score he gives the teacher, and they give no basis for the teacher to analyze his weaknesses and improve. If ratings are a supervisory device, the latter two criticisms are probably valid, but when ratings are used for research, these criticisms are not relevant as there is no attempt made to help the individual teacher. The tendency to be unreliable and the difficulty in determining reliability, however, are important criticisms of subjective ratings.

Thayer supported the viewpoint that a global impression or a more subjective evaluation of teacher effectiveness was superior when he stated:

. . . We now make bold to suggest general impression as the most effective measure of a teacher's worth. (21, p. 377)

.

. . . This general impression is a generalization based upon intimate contacts with a teacher's performances. . . . When a supervisor works intensively with his teachers in the solution of their problems and in the undertaking of new ventures, he will know their strength and their weaknesses more accurately than is possible from a few observations with a score card. (21, p. 378)

Rose summarized the difficulties encountered with ratings as being: problem of untrained raters, halo effect, tendency to judge high or low and the reluctance to use extremes, problem of lack of reliability between raters, problem of weighting, effect on rater of the order of the teachers to be rated, the usual dependence on memory when giving a rating, and the fact that the rater is performing two operations at one time--recording and judging (14, p. 233).

III. PERCEPTION OF BEHAVIOR AND JUDGMENT OF EFFECTIVENESS

Since perception and judgment were essential processes in operation in this research, it was considered necessary to describe some of the literature pertinent to these processes.

Both the description of behavior and the rating of effectiveness imply that the judges perceive the teacher within a certain frame of reference that they have developed. Sherif and Cantril state this very clearly in the following quotation:

Irrespective of the particular way in which an individual acquires an attitude in social life, the literature of social psychology is rife with data which support the formulations reached from our survey of general psychology: that perception and judgment are selective and occur within a referential framework, that frames or points of

reference are inevitably established if an individual repeatedly faces the same stimulus situation, that these frames and points of reference are by no means always confined to consciously accepted instruction or imposed norms but can become established without the individual's realization of it, and that once established these frames and points of reference serve as anchorages for perception and judgment. (19, p. 60)

From a review of many studies Scott notes the following four generalizations about perceptions:

1. That which an individual perceives when confronted with a stimulus situation may not correspond to "objective reality." . . .

2. Several individuals confronted with the same stimulus situation may perceive (i. e., structure and interpret) the same situation quite differently. . . .

3. The manner in which an individual perceives a stimulus situation is affected biological, psychological, social, and cultural factors. . . .

4. The individual responds to his environment in terms of his perceptions of that environment. . . . (17, p. 8, 9)

With regard to the ability to judge others, Taft provided evidence that this may be both a general ability, and a specific ability. The latter ability would vary with the type of judgment to be made and with the situation. The main attributes of judging others seemed to lie in three areas: possessing appropriate judgmental norms, judging ability, and motivation. The development of appropriate

norms is dependent upon the similarity of background of the judge and judged, and also on the proximity of the judge to the judged in regard to their working relations (20, p. 20). Taft believed that his investigation showed that motivation was probably the most important area affecting judgments.

IV. SUMMARY

The literature reviewed described some of the approaches that have been used to study teacher effectiveness. The area of teacher behavior seemed to have been neglected and appeared to be the most fruitful area for research on this topic.

Many studies have been reported on rating of teacher effectiveness. A summary of some of the criticisms of the use of ratings indicated that global impressions by judges in close contact with the teacher might be useful, and that there are many difficulties in using ratings of teacher effectiveness.

Perception of behavior is determined by the frame of reference that the judge has developed and judgments of others appeared to be related to judging ability, development of appropriate norms, and motivation of the judge.

REFERENCES

- (1) Ackerman, W. I. "Teacher Competence and Pupil Change," Harvard Educational Review, 24:273-289, 1954.
- (2) Beecher, D. E. The Evaluation of Teaching. Syracuse: Syracuse University Press, 1949.
- (3) Biber, B. "Problems of Values and Measures in Evaluation of Teaching," Educational Leadership, 15:213-217, January, 1958.
- (4) Gage, N. L., and J. S. Orleans. "Guiding Principles in the Study of Teacher Effectiveness," Journal of Teacher Education, 3:294-298, December, 1952.
- (5) Gray, R. "Teacher Evaluation," Canadian Education, 10:171, June, 1955.
- (6) Hughes, M. M. "Whither Evaluation?" Educational Leadership, 15:208-212, January, 1958.
- (7) Levin, H. "A New Perspective on Teacher Competence Research," Harvard Educational Review, 24:98-105, 1954.
- (8) Melby, E. O. "Role of Evaluation in Improving Teaching," Educational Leadership, 15:218-221, January, 1958.
- (9) Mitzel, H. E. "A Behavioral Approach to the Assessment of Teacher Effectiveness," Monograph. New York: Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, February, 1957. (mimeographed).
- (10) Morris, E. H. Personal Traits and Success in Teaching. Teachers College Contributions to Education, No. 342. New York: Bureau of Publications, Teachers College, Columbia University, 1929.
- (11) Orleans, J. S., et al. "Some Preliminary Thoughts on the Criteria of Teacher Effectiveness," Journal of Educational Research, 45:641-648, May, 1952.
- (12) Payne, E. G. (ed.). Readings in Educational Sociology, Vol. II. New York: Prentice-Hall Inc., 1940.

- (13) Reavis, W. C., and D. H. Cooper. Evaluation of Teacher Merit in City School Systems. Chicago: University of Chicago Press, 1945.
- (14) Rose, G. "Toward the Evaluation of Teaching," Educational Leadership, 15:231-238, January, 1958.
- (15) Rugg, H. "Is the Rating of Human Character Practicable?" Journal of Educational Psychology, 12:425-438, November, 1921; 12:485-501, December, 1921; 13:30-42, January, 1922.
- (16) Ryans, D. G. (chairman), et al. "Report of the Committee on the Criteria of Teacher Effectiveness," Review of Educational Research, 22:238-263, June, 1952.
- (17) Scott, E. L. Leadership and Perceptions of Organization. Research Monograph Number 82. Columbus: The Bureau of Business Research, The Ohio State University, 1956.
- (18) Shane, H. G., and W. A. Yauch. Creative School Administration. New York: Henry Holt and Co., 1954.
- (19) Sherif, M., and H. Cantril. The Psychology of Ego-Involvements. New York: John Wiley and Sons, 1947.
- (20) Taft, R. "The Ability to Judge People," Psychological Bulletin, 52:1-23, January, 1955.
- (21) Thayer, V. T. "Teacher Rating in the Secondary School," Educational Administration and Supervision, 12:361-378, September, 1926.
- (22) Young, L. P. The Administration of Merit-Type Teacher Salary Schedules. New York: Teachers College, Columbia University, 1933.

CHAPTER III

THEORY AND METHODOLOGY

I. FRAMEWORK TO STUDY TEACHER EFFECTIVENESS

A research pattern that includes the important variables which are related to teacher effectiveness will serve as a framework within which theories can be hypothesized and tested. Levin believes there is lack of a conceptual framework for research on this subject. He states that a framework provides an operational philosophy (8, p. 101). Kinney agrees with this when he says that we need to, ". . . Provide a conceptual framework to aid in setting up hypotheses about teacher behavior and related variables." (7, p. 3)

For the purpose of this study, the framework suggested by Mitzel (9, p. 1) was used. He postulated that there are four types of variables that are related to teacher effectiveness.

Type I variables are characteristics of teachers. These are variables concerned with the individual's personality, his background and his education. They differ for almost every teacher, but can be hypothesized as having some relation to teacher effectiveness. They are called prediction sources.

Type II variables are factors from the environment of the school situation. They influence the complex behavior

patterns that are part of the educational process. These situational variables would include environmental factors such as the community, and factors within the school such as the school plant and school organization. These variables would also include pupil factors such as attitudes, interests and abilities. The pupil factors are situational factors that affect teacher and pupil behavior just as teacher variables, Type I, affect teacher and pupil behavior.

Type III variables deal with behaviors. These behaviors of the teachers and pupils are expressions of the individuals in action within a certain situation. Teacher behavior is found in the classroom, in the school, in the community and in the profession. The focal point is the interaction of teacher-pupil behavior in the classroom. Mitzel states, "Considering that classroom behaviors bear such heavy responsibilities in determining educational outcomes, remarkably little is known about them or their effects." (9, p.1)

Type IV variables are the standards with which the results of education are judged. The term intermediate educational goals is used to distinguish them from ultimate goals found in the community. Such factors as pupil growth in subject matter knowledge, in appreciations, and in social skills are some of these variables.

Figure I gives a diagram of this pattern. Solid lines indicate direct effects, and broken lines indicate indirect effects, although the distinction between direct and indirect might be questionable.

In this pattern, teacher variables and pupil variables are direct determinants of teacher and pupil behavior. Environmental variables indirectly influence both. The crucial area is the interaction of teacher and pupil behavior.

Levin states that to have a complete picture of teacher effectiveness one would predict from the antecedent conditions to the classroom behavior and in turn from behavior to classroom effects (8, p. 102).

II. TEACHER BEHAVIOR

From Chapter II and the above section the crucial area in the study of teacher effectiveness appeared to be teacher behavior. The initial step in this study was to determine dimensions of teacher behavior which would be to some extent independent of each other, and would show promise of being related to teacher effectiveness.

Halpin (4, p. 4) suggested that the two dimensions of Consideration and Initiating Structure, which have been found to be valid measures of leader behavior, might be characteristic of teacher behavior in relation to pupils.

Teacher Variables

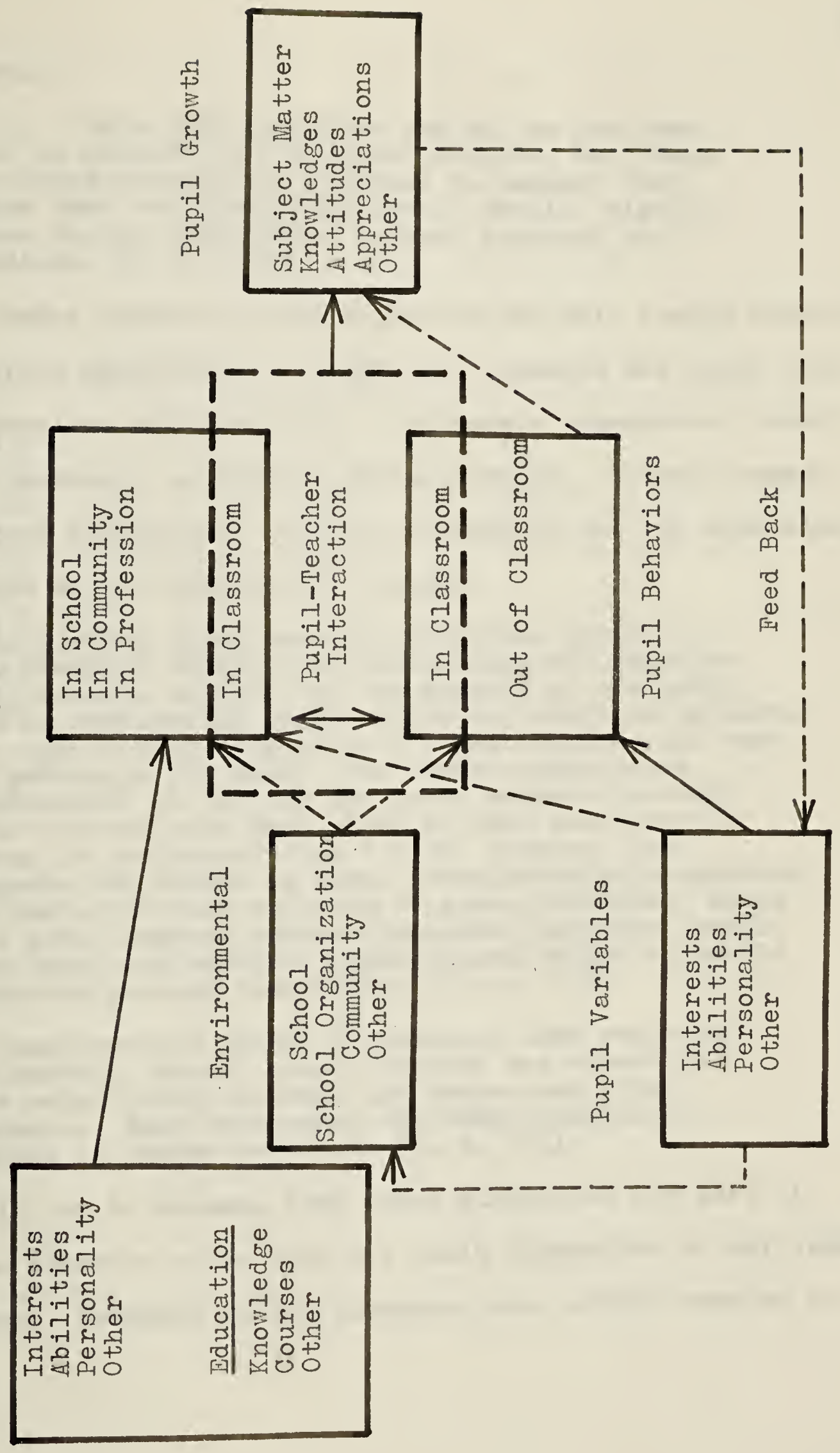


Figure 1 Framework for Research in Teacher Effectiveness

TABLE I. ANALYSIS OF THE

TABLE II. ANALYSIS OF THE

TABLE III. ANALYSIS OF THE

TABLE IV. ANALYSIS OF THE

TABLE V. ANALYSIS OF THE

TABLE VI. ANALYSIS OF THE

TABLE VII. ANALYSIS OF THE

TABLE VIII. ANALYSIS OF THE

TABLE IX. ANALYSIS OF THE

TABLE X. ANALYSIS OF THE

TABLE XI. ANALYSIS OF THE

TABLE XII. ANALYSIS OF THE

TABLE XIII. ANALYSIS OF THE

TABLE XIV. ANALYSIS OF THE

TABLE XV. ANALYSIS OF THE

TABLE XVI. ANALYSIS OF THE

TABLE XVII. ANALYSIS OF THE

TABLE XVIII. ANALYSIS OF THE

TABLE XIX. ANALYSIS OF THE

TABLE XX. ANALYSIS OF THE

TABLE XXI. ANALYSIS OF THE

TABLE XXII. ANALYSIS OF THE

TABLE XXIII. ANALYSIS OF THE

TABLE XXIV. ANALYSIS OF THE

TABLE XXV. ANALYSIS OF THE

He states,

. . . This LBDQ technique has not as yet been used in studies of classroom teachers, but there is strong presumptive evidence to suggest that these same two dimensions are of special significance in the interaction between teachers and students. (4, p. 175)

Leader behavior is described on the Ohio Leader Behavior Description Questionnaire (LBDQ) and research has shown that the dimensions are applicable to aircraft commanders, school superintendents, university administrators, factory foremen and school principals. The definitions of the two dimensions described on the LBDQ are as follows:

Initiating Structure-in-Interaction refers to the leader's behavior in delineating the relationship between himself and the members of his group, and in endeavoring to establish well-defined patterns of organization, channels of communication, and ways of getting a job done. The leader establishes "Structure" in the way the group members interact with him and with each other so that whenever the group is confronted with a novel problem, the members can resort to these Structures-in-Interaction to facilitate the solution of group problems. Hence the group members are not dependent upon the leader for fresh and specific instructions on how to handle each new problem that arises. (4, p. 174)

Consideration refers to behavior that reflects friendship, mutual trust, respect and warmth in the relationship between the leader and group members. This represents the human relations aspect of leader behavior. (4, p. 175)

It can be assumed that these dimensions are part of teacher behavior since they are valid dimensions of any leader behavior. Teachers in the classroom are closely equated to

leaders of a productive group, where goal accomplishment or getting a job done, and human relations are important aims of this leader.

III. JUDGES OF TEACHER LEADER BEHAVIOR AND TEACHER EFFECTIVENESS

The LBDQ provided an instrument through which general leader behavior could be described.

Although the dimensions of Consideration and Initiating Structure might be characteristic of teacher behavior, it was essential that some indication of the extent of their relationship to effective teaching be obtained. This required criteria with which to judge teacher effectiveness and the most common criteria are the judgments of superiors. The Encyclopedia of Educational Research states:

. . . The criterion [of teaching success] most frequently used in studies concerned with pre-service selection of teachers is the judgment of persons in the field, such as superintendents, principals, departmental supervisors, and teaching colleagues. . . .

.

Authorities are in quite general agreement that the judgment of experts is the best available criterion of teaching success. . . . (10, p. 1391)

The selection of judges to rate teaching effectiveness was based on several factors. Firstly, they should have

been in close contact with the teacher over a period of time. Secondly, they should include viewpoints that would give as complete a picture of the teacher as possible. The numerous judges would include the opinions of those who had direct contact with the teacher in the process of teaching. Difficulties in judging effectiveness have been noted in Chapter II, but some of the criticism was regarding the use of such evaluative judgments for purposes other than research. The criterion judge is in no way affecting the future of a teacher or affecting his personal relations with the teacher when he makes a judgment that is to be used solely for research, and is therefore more likely to indicate his true opinion.

To describe the teacher's behavior students would need to be included. On the LBDQ the main description of leader behavior was provided by subordinates, and students are the only ones filling this role in relation to teachers. Levin states,

. . . [from work done at Harvard] the evidence is clear that there is agreement among pupils in a classroom as to how the teacher acts. Methodologically, this gives an important tool for the measurement of teacher behavior. In addition, there is some evidence that principals and pupils tend to confirm each others judgments about the teacher. (8, p. 104)

Ryans says that the teacher's demonstration of effectiveness (his effective behavior) may be measured by means

of objective tests and the judgments of various kinds of persons such as the teacher himself, fellow-teachers, pupils, superintendents, principals, supervisors or lay persons (11, p. 258).

To describe and rate the teacher it was believed that students, other teachers on the same staff, principals, superintendents and the teachers themselves would include all those persons who have direct contact with the teacher. The superintendent would have had the most experience in evaluating teachers and the students would have the least specific ability to judge the teacher's effectiveness as Taft (12, p. 20) has shown that age in children is directly related to the ability to judge others.

Students would have had the most contact with the teacher and thus would be more familiar with the teacher's behavior while superintendents would be less familiar with the teacher's everyday behavior.

IV. HYPOTHESES

The following were the three major hypotheses which resulted from the theoretical design discussed in this chapter.

1. Students and superiors (principals and superintendents) rate those teachers highest in teaching

effectiveness who score above the median on both of the leader behavior dimensions, Consideration and Initiating Structure. Students and superiors rate those teachers lowest in teaching effectiveness who score below the median on both of the leader behavior dimensions.

2. Students tend to relate effectiveness of teaching with teachers' scores on both Consideration and Initiating Structure, but Consideration is more closely related.

3. Superiors tend to relate effectiveness of teaching with teachers' scores on both Consideration and Initiating Structure, but Initiating Structure is more closely related.

V. PROCEDURE FOR TESTING HYPOTHESES

In general the procedure to test the hypotheses was to describe leader behavior and rate effectiveness of teachers through the use of two instruments. These two instruments were to be administered to students, other teachers on the same staff, principals, superintendents, and the teachers themselves in schools in Alberta. This section outlines the instruments used and the design of the research to test the hypotheses.

Instruments

Both of the instruments used were adaptations of other instruments. The instruments were identical for all sets of raters, but five colours were used to facilitate analysis. Both instruments were printed on the same four page booklet, with the behavior description being completed first, followed by the rating scale which was on the back page.

Global rating scale of teacher effectiveness. This very short instrument (Appendix A, p. 124) was adapted from Andrews (1, p. 167).

The instrument gives a global rating. This means that a single rating gives an overall picture or general impression of the individual's opinion of the effectiveness of the teacher.

The rating was applied only to the teaching of students and the directions mentioned not to include other aspects of the teacher's work. This complies with the limitation of effectiveness stated in Chapter I. The directions state, "Please check the statement which in your judgment best indicates how well this teacher does in teaching his students." (Appendix A, p. 124)

The rater was to use as his standard the average teacher of the particular school in which he was at the time of this administration. This was accomplished by adding

the words "for this school" to each of the six categories. This was an attempt to reduce the variability among raters and to avoid the concept of the "ideal" teacher.

The ratings ranged from a possible "1. Outstanding for this school" to "6. Very low for this school." Wandt and Ostreicher (13, p. 6) quote Symonds as believing that seven is the optimal number of scale units for ratings. This scale has six since the opportunity to choose a middle or average rating has been eliminated.

This instrument was used as a validity criterion for the other instrument, the behavior description questionnaire, and as such its own validity is not measured.

The reliability of the rating scale presented a problem. With this instrument reliability could be measured only by the test-retest method. This would require that the names of the raters be recorded on the instruments so that the retest scores could be compared. To do this would destroy the anonymity of the raters and the investigator believed this would change the motivation, and attitude of the raters, to such a degree that it would make the scale much less useful. The only slight indication of reliability would be the consistency from rater to rater for the same teacher and this is discussed in Chapter V, but this gave no indication of the reliability for the same raters and

is considered a definite weakness in the research.

The Ohio Leader Behavior Description Questionnaire (LBDQ). The instrument used in this study to describe behavior was an adaptation of the LBDQ¹.

The LBDQ provided a method whereby the designated leaders of a formal organization might be described by the group members. It specifically states in the directions that it is a description of the leader's behavior and does not ask the respondent to judge whether the behavior is desirable or undesirable.

Halpin states:

The LBDQ was developed by the staff of the Personnel Research Board, The Ohio State University, as one project of the Ohio State Leadership Studies, directed by Dr. Carroll L. Shartle. Hemphill and Coons constructed the original form of the questionnaire; and Halpin and Winer, in reporting the development of an Air Force adaptation of the instrument, identified Initiating Structure and Consideration as two fundamental dimensions of leader behavior. . . . Initiating Structure and Consideration accounted for approximately 34 and 50 percent of the common variance. (3, p. 1)

In a later study using the LBDQ, Halpin concluded that both dimensions were related to aircraft commanders' effectiveness when rated by crews and superiors (5, p. 18).

¹Permission to adapt was obtained from R. M. Stogdill, Ohio State University, February, 1959.

A considerable amount of research was reported regarding the independence of the dimensions. Hemphill reported correlations between the dimensions of .18 for teachers' descriptions of superintendents, .61 for board members' descriptions of the same superintendent, .56 for descriptions of industrial foremen, and .45 in a military situation (6, pp. 92, 93). Halpin reported correlations between the dimensions of .38 and .51 for descriptions of aircraft commanders. He stated that the probability was very slight that a leader would be described as extremely low or extremely high on both dimensions (5, pp. 2, 17).

The LBDQ manual provides no norms, but there are mean scores which were obtained for aircraft commanders and school superintendents (3, p. 10).

In using the instrument the respondent describes the leader on forty items by indicating the frequency with which he engages in each behavior by circling a letter representing one of five adverbs, "always," "often," "occasionally," "seldom," and "never." It is scored by assigning values of 4, 3, 2, 1 and 0 to "always," "often," "occasionally," "seldom," and "never," respectively. Of the forty items, ten are not scored so that each dimension contains fifteen items.

In administering the questionnaire, Initiating

Structure and Consideration are not mentioned. With regard to the number of respondents per leader, Halpin states,

Experience suggests that a minimum of four respondents per leader is desirable and that additional respondents beyond ten do not increase significantly the stability of the index scores.
(3, p. 7)

The Teacher Leader Behavior Description Questionnaire (TLBDQ) (Appendix A). The administration, the scoring and the purpose of the adapted instrument were similar to that of the LBDQ described above. The original questionnaire was changed to make it more suitable to a classroom situation.

The ten unscored items were discarded. Thirteen of the fifteen Initiating Structure items remained the same as did nine of the fifteen Consideration items. This is with the exception of the minor changes for the use of "group" to "class" and "group members" to "students." Of the remaining eight items, some changes were made, but with six of them the investigator tried to keep the intent of the original item.

With two items considerable change was made. These were items number 17 and 19 which were both part of the Consideration dimension. The items used were chosen from a list of behaviors reported by Beecher (2, pp. 64-78) as being considered important by students. Since these items were so different from the originals, it was considered necessary to

establish the amount of agreement with the total remaining Consideration score. Item number 17 gave a correlation of .48 (significant at the .01 level, $N = 200$) with the total Consideration score and item number 19 correlated .38 (significant at the .01 level, $N = 200$). These were considered sufficiently high to include them as part of the questionnaire. These results were obtained from the study proper.

To establish the reliability a random sample of 100 questionnaires was selected from the 989 questionnaires that were satisfactorily completed in the study. This sample was so chosen that the ratio of students, to other teachers, to principals, to superintendents, to teachers themselves was the same as in the total group. This sample of 100 was used to check the reliability. Using the Spearman-Brown formula the corrected split-half reliabilities were .75 and .86 for Initiating Structure and Consideration, respectively. Halpin (3, p. 6) reported .83 and .92 for the same dimensions. Initiating Structure seems to have a lower reliability. The investigator considered the reliability for the Initiating Structure dimension on the TLBDQ as lower than desirable.

The Pilot Study

To examine the wording of the items for use with students a pilot study was conducted in a junior high school and a senior high school in Edmonton. This involved nine students, two teachers and a principal in each of the two schools. There was no attempt made to establish the validity of the instrument in relation to effectiveness as it was felt that this was one of the purposes of the study proper.

Only a few minor changes were made in the wording. It was found that students had difficulty with the three negative items which were in the original LBDQ. The revision left only one negative item and here the values assigned were reversed from the usual method.

The investigator found that senior high school students expressed no difficulty in interpreting and understanding the items. Some of the teachers thought that a few items were ambiguous.

In the junior high school there was more difficulty expressed by the students and interrogation by the investigator indicated some variation in the interpretation of certain items. This was especially true with grade seven. For this reason the original intent to include grade seven in the study was dropped.

It was difficult to establish objectively, but the investigator felt that the descriptions given by junior high school students were less valid due to misunderstanding of the terminology used in the items.

Plan for Collecting the Data

To test the hypotheses the two instruments were to be administered to the five sets of respondents. Thirty-six provincially appointed superintendents were contacted by letter (Appendix B) asking for the cooperation of one school in their area. The schools were required to have at least ten teachers on the staff and at least seven teaching above grade seven. The investigator asked that no school be included in the research unless every member on the staff wished to take part.

The plan was to use non-city schools as it was believed that the relationship between the superintendent and teachers might not be similar to the relationship between the superintendent and teachers in divisions. Elementary teachers were not to be included among those to be described or rated as elementary pupils would not be able to understand the wording of the TLBDQ.

The plan was to obtain the cooperation of twelve schools. In each school five teachers, to be the centre of

the study, would be chosen at random from those teaching above grade seven. At the same time five other teachers were to be chosen to describe the behavior and rate the effectiveness of the five teachers first chosen. For each of the five teachers that were the centre of the study, ten students would be chosen at random from all the students taught by that teacher. The teachers themselves, as well as their respective principals and superintendents, would also describe and rate the teachers.

This plan would involve 60 teachers as the centre of the study who would be described and rated by 600 students, by 60 other teachers, by 12 principals, by 12 superintendents and by the teachers themselves.

VI. SUMMARY

There is need for a framework within which to study teacher effectiveness. A pattern was suggested in which the crucial area is that of teacher and pupil behaviors. In this study the area of teacher behavior is investigated and it was suggested that the two dimensions of teacher behavior which might be related to teacher effectiveness are the leader behavior dimensions of Consideration and Initiating Structure. The types of judges to provide the descriptions of behavior and the ratings of effectiveness were discussed

and the specific hypotheses to be tested were stated. The procedure to be used in testing the stated hypotheses was outlined by describing the two instruments and the plans for collecting data.

REFERENCES

- (1) Andrews, J. H. M. "Administrative Significance of Teacher Differences," Unpublished Doctoral dissertation, The University of Chicago, 1957.
- (2) Beecher, D. E. The Evaluation of Teaching. Syracuse: Syracuse University Press, 1949.
- (3) Halpin, A. W. Manual for the Leader Behavior Description Questionnaire. Columbus: Bureau of Business Research, The Ohio State University, 1957.
- (4) Halpin, A. W. "The Behavior of Leaders," Educational Leadership, 14:172-176, December, 1956.
- (5) Halpin, A. W. Studies in Aircrew Composition. III. The Combat Leader Behavior of B-29 Aircraft Commanders. Bolling Air Force Base, Washington, D. C.: Human Factors Operations Research Laboratories, 1953, (HFORL Memorandum No. TN-54-7).
- (6) Hemphill, J. K. "Leader Behavior Associated with the Administrative Reputations of College Departments," Leader Behavior; Its Description and Measurement. R. M. Stogdill and A. E. Coons, editors. Pp. 74-102. Research Monograph Number 88. Columbus, Ohio: Bureau of Business Research, The Ohio State University, 1957.
- (7) Kinney, L. B. "What Is a Good Teacher?" School of Education Bulletin, Utah State Agricultural College, Utah, 1954.
- (8) Levin, H. "A New Perspective on Teacher Competence Research," Harvard Educational Review, 24:98-105, 1954.
- (9) Mitzel, H. E. A Behavioral Approach to the Assessment of Teacher Effectiveness, Monograph. New York: Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, February, 1957. (mimeographed).
- (10) Monroe, W. S. (ed.). Encyclopedia of Educational Research. New York: The MacMillan Co., 1952.

- (11) Ryans, D. G. (chairman), et al. "Report of the Committee on the Criteria of Teacher Effectiveness." Review of Educational Research, 22:238-263, June, 1952.
- (12) Taft, R. "The Ability to Judge People," Psychological Bulletin, 52:1-23, January, 1955.
- (13) Wandt, E., and L. M. Ostreicher. Variability in Observed Classroom Behaviors of Junior High School Teachers and Classes. Research Monograph Number 16. New York: Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, 1953.

CHAPTER IV

COLLECTION OF THE DATA

The data for this research were collected from twelve schools in the Province of Alberta. The staffs of the twelve schools volunteered for the project. The investigator visited each school in April and May of 1959.

I. METHOD OF COLLECTING THE DATA

The Teachers Who Were the Centre of the Study

At each school five teachers were chosen from a hat on the day of the visit. There were two exceptions to this procedure. In one school there were only four teachers teaching above grade seven; thus the entire four were used. In another school one teacher volunteered to change places with a teacher who already had been selected. No principals were included as teachers. A total of fifty-nine teachers were chosen as described above and constituted the core of the study.

The five teachers received a blue questionnaire to describe and rate themselves. They completed the questionnaires during their spare time and returned them to the investigator during the day.

Whenever the word "teacher" or the words "teachers

themselves" are used they refer to this group of fifty-nine teachers who were the centre of the study.

The Other Teachers on the Same Staff

At the same time as the above five teachers were selected, five other teachers were chosen at random from the staff. This peer group was used to describe and rate the original five teachers. They are referred to as the "other teachers" in the remainder of the study.

Each other teacher received five green questionnaires and completed them individually during the day. They completed one questionnaire on each of the five teachers.

The number of other teachers varied from two to five. In one school two were all that were available. In two schools three were used. In four schools four other teachers were utilized and in the remaining five schools five other teachers from each school completed the questionnaire. The average was over four other teachers per teacher. Questionnaires completed by the 50 other teachers totalled 241. In all but one school the other teachers were from senior or junior high school.

The Students

The plan was to chose ten students to describe and

rate each of the five teachers in the school. The selection of the students was carried out with the aid of the principal.

For each teacher an alphabetical list was obtained of all the students that he or she taught. The ten students chosen were as equally distributed as possible throughout all the grades taught by the teacher while at the same time giving some weight to the number of pupils in each grade or class. For example, if a teacher taught grades ten, eleven, and twelve with more students in grade ten, four students were selected from grade ten and three from each of the other two grades. The students were chosen from alphabetically arranged lists using numbers predetermined by the investigator. In all cases the students involved were taught one or more subjects by the teacher they described. Some students had been taught by the specific teacher for a number of years and some for just the present year. In no case did a student describe more than one teacher; once chosen for a specific teacher, they were not chosen to describe another teacher.

There was one exception to the above random selection of students. In one class the teacher had a disproportionate number of repeaters compared to the normal class. A division was made between the repeaters and non-repeaters and the students were chosen at random from each group. The resulting group of students had fewer repeaters than if the

selection had been made at random from the whole class. In a few instances the principal provided an additional student when one of those previously selected was not available.

The smallest number of students used was 7 per teacher. The total number of students was 571 which averages over 9.5 students per teacher. Of this total 222 were from grades eight and nine and the remaining 349 were from senior high school.

The students gathered in a room and the white questionnaire was administered separately to each group by the investigator. The directions (Appendix A) emphasized that the teachers were chosen by chance and that the results were for research. Each student completed the questionnaire on his own and returned it to the investigator before leaving the room.

The Principals

During the day of the visit the principal completed one yellow questionnaire on each of the five teachers. Thus, there were fifty-nine questionnaires completed by twelve principals.

The Superintendents

The superintendents, all provincially appointed,

completed one pink questionnaire on each of the five teachers from the one school in their area. In all but three cases they mailed them to the investigator. There were fifty-nine questionnaires completed by twelve superintendents.

II. THE SCHOOLS AND TEACHERS INVOLVED

The hypotheses of this study involved a belief that elements of teacher leader behavior are common to teachers anywhere and under all circumstances. Thus a discussion of the breadth of coverage is considered helpful at this time. A brief glance at the schools and teachers involved showed that the coverage was very broad.

City schools were not used. Elementary teachers were omitted because it would not be possible to have elementary students answer the questionnaire.

Geographically the schools covered many parts of the province, but in the main were centred around Edmonton. No two schools were in the same inspectorate. All schools, but one, were in divisions and counties. No separate school districts were included.

The fifty-nine teachers studied covered a wide range of subjects taught, experience, training and background. There were twenty-one females and thirty-eight males. Subjects taught included all academic courses plus home

economics, commercial, physical education, art, music, and industrial arts. Some of the teachers were travelling specialists. They taught all grades from eight to twelve. Some taught only one grade while others taught four grades. They ranged from one to seven years of academic and professional training beyond high school. In total experience the range was from two to thirty-five years and the experience in their present school ranged from one to seventeen years. Some of the teachers were recently from England, The United States, Iceland, and Saskatchewan. The size of the schools ranged from a low of eleven members on a staff to a high of twenty-seven.

III. SUMMARY

The data for this research were collected from 12 schools in Alberta through personal visitation by the investigator. It involved 59 teachers who were the centre of the study and whose behaviors were described and effectiveness rated by themselves, by 571 students, by 12 principals, by 12 superintendents, and by 50 other teachers completing 241 questionnaires.

Questionnaires were administered to 1,004 persons. Fifteen were eliminated because they were only partially completed or the directions were not followed. The

remaining 989 questionnaires were utilized in the study.

The teachers involved in the study included a wide variety of subject specialists, both academic and non-academic; they were teachers with a wide range of years of training, total experience, and experience in their present school. Some were recently from areas outside Alberta.

CHAPTER V

RESULTS, ANALYSIS AND DISCUSSION OF EFFECTIVENESS RATINGS

I. RESULTS AND ANALYSIS

Group Averages and Deviations

In analysing the effectiveness rating scores for the fifty-nine teachers there was no assumption of normal distribution of scores nor of equal intervals on the rating scale. Thus, in most cases non-parametric statistics were used.

Table I summarizes the medians, quartile points, and quartile deviations for the five sets of raters when judging the effectiveness of the fifty-nine teachers. Since a rating of "1" was "outstanding for this school" and a rating of "6" was "very low for this school" (see Appendix A), a low score indicated a higher rating and a high score indicated a lower rating.

In the case of the students and other teachers, the score for each teacher was derived from the median of the ratings. In the case of the other three sets of raters, the principals, the superintendents, and the teachers themselves, single ratings provided their judgment about each teacher's effectiveness. Thus, only the first two sets of ratings utilized the opinion of more than one individual per teacher.

TABLE I

MEDIANS, QUARTILE POINTS AND QUARTILE DEVIATIONS OF THE
EFFECTIVENESS RATINGS FOR FIFTY-NINE TEACHERS

=====					
Ratings provided by:					
	Students	Other Teachers	Principals	Superintendents	Teachers
Q ₁	1.91	1.87	1.94	2.09	2.35
Q ₂	2.39	2.38	2.48	2.78	3.04
Q ₃	2.99	3.06	3.39	3.42	3.43
Q	.54	.60	.72	.66	.54
=====					

NOTE: If the teachers were average for the school, the median should be 3.50.

Higher ratings are indicated by smaller scores and lower ratings are indicated by higher scores. See global rating scale Appendix A.

Whenever the effectiveness ratings by students or other teachers are discussed, the ratings refer to the single, composite rating obtained for each teacher from the median of the ratings by the ten students and from the median of the ratings by the five other teachers.

The highest median rating in Table I was provided by students and other teachers, followed by principals, superintendents, and the teachers themselves. All five sets of raters saw the average of the fifty-nine teachers as being above the average of the scale which was 3.50. Therefore, the tendency to rate high found in other studies was supported. Contrary to the expectations expressed to the investigator by many teachers, the students rated the teachers highly.

The principals spread their ratings out over a wider range than did the other raters as indicated by the quartile deviations (Q). The fact that ten students and five other teachers had their ratings averaged to provide a single rating would tend to nullify any extremes among these two groups of judges.

For the students' ratings the quartile deviation was calculated for each teacher. Of the 59 deviations so calculated the greatest was .91 and the lowest was .25 with the mean at .55. For the other teachers' ratings the quartile

deviations were calculated for 54 of the 59 teachers and they ranged from 1.00 to 0 with the mean at .48. There was no significant difference at the .05 level between the mean of the deviations for the students and the mean of the deviations for the other teachers.

An analysis of the range of scores was provided by the frequency with which the extremes of the rating scale were used. Table II summarizes this information. Students and teachers themselves did not rate any teacher in the two lowest categories. The number of teachers who were rated "1. Outstanding for this school" were few, but fairly consistent for four of the sets of raters.

For use in a later part of the study the ten students for each teacher were divided into two groups at random. The purpose of this procedure was to provide independent judges for effectiveness ratings and behavior descriptions. Thus, one half-group of students would provide effectiveness ratings and the other half-group of students would provide behavior descriptions. At this stage of the analysis only the half-group which provided the effectiveness ratings was studied. The median of the effectiveness ratings of this half-group of students was 2.38 which was practically identical to the median obtained from the whole group of students (2.39). The quartile deviation was .60 which was slightly higher than that reported for the whole group (.54).

TABLE II
FREQUENCY OF EXTREME SCORES USED IN RATING FIFTY-NINE TEACHERS

Extreme scores	Frequency of ratings provided by:			
	Students	Other teachers	Principals	Superintendents Teachers
1	2	4	3	3 0
2	31	29	27	20 9
5	0	1	5	5 0
6	0	0	1	0 0
Total number of teachers rated at the extremes	33	34	36	28 9

Consensus

Table III indicated the rank-order correlations between effectiveness ratings provided by the sets of raters. There was a large number of ties in the rank-orders, thus, the formula given by Siegel (1, p. 203) for tied ranks was used.

There was a consistent lack of relationship between self-ratings and effectiveness ratings by the other sets of judges.

The superintendents and principals showed the most agreement and the superintendents and students showed the least agreement (excluding the self-ratings). To test the significance of the difference between these two correlations parametric statistics had to be used. Since N was 59 the frequency was not too small to use parametric statistics, but the procedure can be criticized on the bases that the intervals of the scale were not equal and the scores might not be normally distributed. The z' transformation was applied and the difference between the correlations of .68 and .41 was found to be significant at the .05 level. Thus, one may conclude, keeping in mind the above criticism, that there was significantly more agreement between superintendents and principals than there was between the superintendents and students on the effectiveness ratings. Excluding self-ratings,

TABLE III
RANK-ORDER CORRELATIONS BETWEEN RATERS

	Students	Other teachers	Principals	Superintendents
Other teachers	.56**			
Principals	.43**	.53**		
Superintendents	.41**	.52**	.68**	
Teachers	-.02	.02	-.01	.05

** significant at the .01 level.

none of the other differences between correlations were significant. Except for the highest and lowest correlations, which were different, the correlations and therefore the agreement between raters appeared to be similar.

A further indication of consensus was derived by calculating Kendall's Coefficient of Concordance (W) (1, p. 229). This coefficient provided the amount of consensus in the rank-order of the effectiveness ratings by more than two judges. It did not indicate the degree of correctness of the rankings, as all judges could agree completely, yet be wrong. It is, however, a useful statistic when there are numerous judges and when there is no external criterion with which to check the validity of the ratings. The range of possible values for W is from zero to plus one. Thus, if W was equal to zero, it would indicate no agreement, while if W was plus one, it would indicate perfect agreement. An indication of the meaning of the W coefficients was obtained by translating them to an equivalent average, rank-order correlation (1, p. 229). This is the average of all the possible correlations between the judges included for any one W coefficient.

Table IV summarizes the W coefficients for three combinations of judges. The first combination included those judges from within the school other than the teacher himself; the second consisted of the superintendent in addition to the

first three, and the last combination included all five sets of judges at once. As an additional judge was added the amount of consensus decreased although the equivalent average, rank-order correlation was highest for the second combination. All W coefficients were significant at the .01 level, but they did not indicate a great deal of consensus.

TABLE IV
CONSENSUS AMONG JUDGES ABOUT EFFECTIVENESS RATINGS

Judges	W*	Equivalent average, rank-order correlation
(1) Students, Other teachers, and Principals	.64	.42
(2) Students, Other teachers, Principals, and Superintendents	.61	.48
(3) Students, Other teachers, Principals, Superintendents, and Teachers themselves	.43	.29

* All W coefficients were significant at the .01 level.

The amount of agreement at the extremes of the rating scale was analyzed. For a rating of "1" there was no teacher so rated by all five sets of judges. One teacher appeared in this category for four sets of raters. No teacher was rated "1" by three sets of raters and only one was so rated by two

sets of raters. Six of the teachers were rated in category "1" by only one set of raters. Thus, out of eight teachers who were rated in category "1" of the rating scale, only two were placed in this category by more than one set of judges. Only one teacher was rated "6" and that was by a principal. There was not a teacher placed in category "5" by all five sets of raters, nor by four or three sets of raters. Two teachers were rated "5" by two sets of raters and seven were placed in this category by only one set of raters. Thus, out of nine teachers who were rated in category "5", only two were placed in this category by more than one set of judges. When compared to the amount of agreement possible, it can be said that there was lack of agreement about rating of teachers at the extremes of the scale.

Variations by Schools from the Average for the Scale

Originally, the investigator considered making an adjustment to the effectiveness ratings to take into account the variation of the average score, of the five teachers on each staff, from the average score of the scale. Since the scale was so worded that the judges were to use as their standard the average teacher of each school, this might have been a logical adjustment. If all the teachers on a staff were rated, the median rating should be close to 3.50 which

was the average of the scale. Due to the fact that only five teachers in any one school were used and those selected were from above grade seven, the assumption that the average rating in each school should be the average of the scale, was dropped. It was considered plausible that the average rating of the five teachers selected might vary from 3.50. An investigation showed that for effectiveness ratings provided by students the medians were nearly all around one unit above the centre of the scale. The consistency of the average rating of the five teachers from school to school was very noticeable. Some of the previous rank-order correlations were calculated with the use of adjusted ratings and there did not appear to be any significant difference from the correlations using the raw ratings.

II. DISCUSSION

The results indicated that the median effectiveness ratings of the fifty-nine teachers were above the average for the scale. The teachers rated themselves as being lower in average effectiveness than did any of the other judges.

There was some agreement among the judges as to the rank-order of the effectiveness ratings. Since only the lowest and highest rank-order correlations were significantly different, there was some similarity in the amount of agree-

ment between pairs of judges. The total consensus although significant in all cases was not considered very great. There appeared to be just about as much consensus with the superintendent added to the other three judges (excluding the teachers themselves) as there was with just these three (students, other teachers, and principals). At the extremes of the rating scale there was lack of agreement compared to complete agreement.

The greatest amount of disagreement was between the teachers' effectiveness ratings of themselves and the effectiveness ratings by the remaining sets of judges. Also, as noted above, the teachers judged themselves as being lower in average effectiveness than did any of the other judges. The investigator felt that motivation rather than frame of reference or ability to judge was the important factor. Many teachers tended to be modest. This was evident not only in the results, but also in their comments to the investigator. Some stated that the scale was very difficult to use since it did not include a category "average for this school." This was the rating that many teachers indicated they would prefer to give themselves. There is also a possibility that competent teachers more readily realize their inadequacies and tend to rate themselves lower than external judges would rate them.

The literature in Chapter II regarding judging seemed to indicate that superintendents would be the most competent judges due to their broader experience allowing them to develop more appropriate norms, and due to more practice in actual evaluation which might develop their ability to judge. It was not possible to verify the correctness of their judgment, but they judged the teachers as a whole as being closer to the mid-point of the scale than did any other external judge. Also, they agreed to a greater extent with the principals than they did with any other group.

The highest relationship, between the superintendents' and the principals' effectiveness ratings, was significantly different from the lowest relationship, between the superintendents' and the students' effectiveness ratings. The greatest amount of agreement between superintendents and principals might be explained by the fact that they were in the most similar position with regard to experience in evaluating teachers and supervisory contact with teachers. The superintendents and students would be in the most dissimilar position since the literature indicated that students would have the least ability to judge other individuals, and would have the least opportunity to develop a frame of reference for judging. Also, the superintendents had the least contact with the teachers in the classroom while the students had the most contact with the teachers in the classroom.

REFERENCES

- (1) Siegel, S. Nonparametric Statistics for the Behavioral Sciences. Toronto: McGraw-Hill Book Co., 1956.

CHAPTER VI

RESULTS, ANALYSIS AND DISCUSSION OF DESCRIPTIONS OF TEACHER LEADER BEHAVIOR

I. RESULTS AND ANALYSIS

Group Averages and Deviations

The scores on the TLBDQ were calculated separately for the dimensions of Consideration and Initiating Structure. A relatively high score on either dimension indicated that the teacher behaved in a certain manner more often than a teacher did with a relatively low score. The averages for the entire group of teachers were derived by calculating the medians.

Table V summarizes the medians, quartile points, and quartile deviations for the five sets of judges when describing the behavior of the fifty-nine teachers.

In the case of the students and other teachers, means were used to give an average measure per teacher. This was the procedure used with the LBDQ, and also, the investigator believed that extremes in behavior scores should be given full weight in computing the average description for each teacher. In the case of the other three sets of judges, the principals, the superintendents, and the teachers themselves, single scores on each dimension provided their judgment about each

TABLE V

MEDIAN, QUARTILE POINTS AND QUARTILE DEVIATIONS OF THE
BEHAVIOR DESCRIPTIONS FOR FIFTY-NINE TEACHERS

	Behavior descriptions provided by:					
	Students	Other teachers	Principals	Superintendents	Teachers	
Q ₁ -I.S. -C.	44.1 43.0	49.6 45.8	48.1 45.9	49.0 48.6	48.7 47.2	
Q ₂ -I.S. -C.	41.4 40.3	46.2 41.9	42.0 40.4	45.2 42.4	42.8 43.6	
Q ₃ -I.S. -C.	38.4 35.4	43.8 38.4	36.8 35.2	35.9 36.4	40.2 38.9	
Q -I.S. -C.	2.8 3.8	2.9 3.7	5.6 5.4	6.6 6.1	4.2 4.2	

NOTE: I.S. - Initiating Structure.
C. - Consideration.

teacher's behavior. Whenever the behavior descriptions by students or other teachers are discussed, they refer to the single, composite score obtained for each teacher from the mean of the scores by the ten students and from the mean of the scores by the five other teachers.

As shown in Table V all the medians for Initiating Structure were higher than those for Consideration except for teachers' descriptions of themselves. This was opposite to the result found by Halpin, with non-teachers, when the medians for superintendents and a sample of aircraft commanders were higher on the Consideration dimension (1, p. 10). There was slightly more agreement on the medians for Consideration than on the medians for Initiating Structure.

The highest Initiating Structure median was provided by other teachers and the lowest by students. The highest Consideration median was provided by teachers' descriptions of themselves and the lowest was provided by students.

The quartile deviations (Q) of the descriptions of the fifty-nine teachers were lowest for the students and other teachers, while the superintendents showed the greatest deviations in their descriptions of the fifty-nine teachers.

The variability of the descriptions provided by the ten students and five other teachers was calculated for each of the fifty-nine teachers by obtaining the standard

deviation. The means of the standard deviations on Initiating Structure were 6.24 for students and 6.03 for other teachers. These were not significantly different at the .05 level. The means of the standard deviations on Consideration were 6.81 for students and 6.09 for other teachers with the difference not being significant at the .05 level. There was no significant difference between the mean standard deviation on Consideration and the mean standard deviation on Initiating Structure for students nor for other teachers.

As was reported in Chapter V (see page 57) the ten students for each teacher were divided into two groups at random. One group was used to provide effectiveness ratings and the other group was used to provide behavior descriptions. An analysis of the half-group of students used to provide descriptions of behavior resulted in an Initiating Structure median of 41.2 and Q of 3.7, while the Consideration median was 39.8 and Q was 3.8. This indicated that the medians of the descriptions provided by one of the half-group of students were very similar to those provided by the whole group of students (Table V). The quartile deviations were the same for the half-group of students and the whole group of students on Consideration, but on Initiating Structure the half-group had a larger deviation than did the whole group.

Independence of the Dimensions

The background research on the LBDQ indicated that, although the two dimensions were related, there was reason to believe they provided scores on dimensions which were somewhat independent. The independence of the two dimensions on the TLBDQ was checked by calculating the product-moment correlations between the dimensions for each of the five sets of raters and the results are shown in Table VI.

TABLE VI

CORRELATIONS BETWEEN INITIATING STRUCTURE AND
CONSIDERATION SCORES FOR EACH SET OF JUDGES

=====	
Judges	Correlations*
Students	.58
Other teachers	.52
Principals	.48
Superintendents	.61
Teachers	.38
=====	

* All correlations were significant
at the .01 level.

The greatest amount of independence was found between the dimensions for the behavior descriptions by the teachers themselves and the least amount of independence for the

behavior descriptions by the superintendents. The results reported in this study were not greatly different from the various correlations reported and accepted for the LBDQ, but the investigator does not consider them indicative of a high degree of independence.

As a further analysis of independence two other correlations were calculated. Both correlations were obtained from the data supplied by the half-groups of students. The first was the correlation between the dimensions for the half-group of students that were used to provide behavior descriptions (see page 71). The correlation was .52 as contrasted with .58 reported for the whole group of students in Table VI. Thus, there was about the same amount of independence between the dimensions when the half-group described the teachers as when the whole group described them. The second correlation was determined by using the Initiating Structure scores from one half-group and the Consideration scores from the other half-group. In this instance the describers were all students, but the scores on one dimension were provided by one group while the scores on the other dimension were provided by the other group. The resulting correlation was .38. Although this was lower than either of the other correlations reported above, the differences were not statistically significant at the .05 level. The difference was tested by making the z'

transformation. This procedure was used throughout to test the significance of the difference of two correlations.

Consensus

The product-moment correlations between the descriptions by the various judges are recorded in Tables VII and VIII and they indicate the amount of agreement between judges. There was some agreement between judges on their descriptions of both dimensions of a teacher's leader behavior with students. As in the results of the effectiveness ratings discussed in Chapter V, there was little or no agreement between behavior descriptions by the teachers themselves and behavior descriptions by other judges. The one noticeable difference between these results and the correlations recorded for the effectiveness ratings was the significant correlations obtained between the students' descriptions and the teachers' descriptions of themselves on both dimensions.

On both dimensions the greatest amount of agreement was between the principals and the other teachers, while the least amount of agreement (excluding self-descriptions) was between the superintendents and the students on Initiating Structure and between the superintendents and the other teachers on Consideration.

For Initiating Structure the difference between the

TABLE VII

CORRELATIONS BETWEEN JUDGES ON THE INITIATING STRUCTURE SCORES

	Students	Other teachers	Principals	Superintendents
Other teachers	.51**			
Principals	.38**	.63**		
Superintendents	.29*	.40**	.48**	
Teachers	.39**	.23	.14	.17

* significant at the .05 level.

** significant at the .01 level.

TABLE VIII

CORRELATIONS BETWEEN JUDGES ON THE CONSIDERATION SCORES

	Students	Other teachers	Principals	Superintendents
Other teachers	.55**			
Principals	.44**	.60**		
Superintendents	.52**	.38**	.43**	
Teachers	.42**	.23	.22	.09

** significant at the .01 level.

correlation showing the most agreement (principals and other teachers) and that showing the least agreement (superintendents and students) was significant at the .05 level. Excluding correlations with self-descriptions, there were no significant differences between other pairs of correlations. One would conclude that the amount of agreement between pairs of judges was similar.

For Consideration the difference between the correlation showing the most agreement (principals and other teachers) and that showing the least agreement (superintendents and other teachers) was not significant at the .05 level. Since the correlations between all pairs of judges (excluding self-descriptions) were not significantly different from each other, one would conclude that the amount of agreement between pairs of judges was similar.

The similarity between correlations of Initiating Structure and the correlations of Consideration was noticeable and could be observed by comparing Tables VII and VIII. There was one exception to this apparent similarity and that was the correlations between superintendents' scores and students' scores. The correlation of Consideration scores appeared to be much higher than that for Initiating Structure scores, but the difference was not significant at the .05 level. All correlations with scores provided by students were higher for

Consideration than they were for Initiating Structure, but the differences were not significant.

On both dimensions the students exhibited some agreement with all other judges including the teachers themselves. All of the eight correlations involving the students were significant at the .01 level except that between the superintendents and the students for Initiating Structure, which was significant at the .05 level. This agreement with all other judges was exhibited only by students.

The Concordance Coefficient (W) was calculated for the rank-order of the scores on both dimensions of behavior. Table IX summarizes these results. An indication of the meaning of the W coefficients was obtained by translating them to equivalent, rank-order correlations (3, P. 229).

All coefficients were significant at the .01 level, but did not indicate a great deal of consensus on what were believed to be easily observed behavior items. There was a slight tendency for more agreement on the Consideration dimension than on the Initiating Structure dimension. The consensus decreased with additional judges. This was also observed with the effectiveness ratings in Chapter V.

TABLE IX
 CONSENSUS AMONG JUDGES ABOUT BEHAVIOR DESCRIPTIONS

Judges	Behavior dimension	
	Initiating	Consideration
	Structure W*	W*
(1) Students, Other teachers, and Principals	.60 (.40)	.69 (.51)
(2) Students, Other teachers, Principals, and Superintendents	.52 (.36)	.61 (.48)
(3) Students, Other teachers, Principals, Superintendents, and Teachers themselves	.45 (.31)	.48 (.35)

* All W coefficients were significant at the .01 level.

NOTE: The equivalent average, rank-order correlations are in brackets.

II. DISCUSSION

The results indicated that all judges except the teachers themselves believed that the teachers on the average exhibited more Initiating Structure behavior than Consideration behavior. The quartile deviations showed that the superintendents deviated more on their descriptions of the fifty-nine teachers on both dimensions than did any other judges. In the mean of the variabilities of the descriptions provided by the ten students for each teacher there was no significant difference between the dimensions. The same result was observed for other teachers. There was as much mean variation on Consideration as there was on Initiating Structure. Also, there was no significant difference between the means of the variabilities of the ten students' descriptions and the means of the five other teachers' descriptions on both dimensions.

As with the effectiveness ratings, the results have shown that there was some agreement between most pairs of judges on the scores of both leader behavior dimensions. The results were not high enough to indicate that the TLBDQ could be used by one or two sets of respondents with the assumption that the behavior descriptions would be all-inclusive. The results indicated that these five sets of raters used varying frames and points of reference in their perception of the leader behavior of the teacher.

The amount of general consensus decreased as more judges were included. There was slightly more consensus on Consideration than on Initiating Structure. The amount of consensus on the effectiveness ratings did not vary greatly from the amount of consensus on the behavior dimensions. Thus, it was just as difficult to agree on the teacher's leader behavior in the classroom as it was to agree on the teachers' effectiveness.

The students agreed to some extent with all of the other four sets of judges on both dimensions. Students may be in a position to have developed a frame of reference which includes parts of the frames of reference of all the other judges. This could be expected since the students would have experienced the teacher's classroom behavior to a much greater extent than any of the other judges.

The highest and lowest correlations for Initiating Structure were significantly different and indicated that principals and other teachers exhibited the most agreement while students and superintendents exhibited the least (excluding the teachers themselves). When everyday behavior of the teacher in the classroom is being described, those who were the closest in similarity of contact with the teacher were the principal and other teachers, since most of the principals in this study spent the majority of their

time teaching. The superintendents and the students were in the most dissimilar position relative to describing the behavior of the teacher and on the Initiating Structure dimension they showed the least agreement.

The highest and lowest correlations for Consideration were not significantly different, when the correlations with self-descriptions were excluded. The amount of agreement between pairs of judges was very similar.

The fact that there was no significant difference between the highest and lowest correlation (excluding self-descriptions) on Consideration, that there tended to be more agreement on the Consideration dimension than on the Initiating Structure dimension, and that two judges, who were remote from each other, (superintendents and students) exhibited a relatively higher agreement, might be partly explained by the nature of the Consideration dimension. The human relations behavior of the teacher in the classroom might be much the same as his relations with other individuals in general. Thus, superintendents, although remote, might more easily infer from their contact with the teacher outside the classroom, what human relations behaviors the teacher would exhibit in the classroom. This does not appear to be as evident for Initiating Structure. The teacher's means of developing particular structure between himself and the

members of the class to attain goals and to provide for a profitable learning situation might not be so easily inferred especially by a person as remote as the superintendent. It should be noted that many of the local, provincially appointed, superintendents do not have much classroom contact with senior high school teachers.

The position of the other teachers in this study deserves discussion. The relatively higher agreement that was exhibited between them and principals on both dimensions than between any other pair of judges was unexpected. One possible reason for this has been explained above. They and the principal are probably in the most similar position relative to the teacher when it comes to describing behavior. Yet the other teachers, rarely, if ever, observe another teacher in the classroom. In some large staffs they have very little association with some members of the staff. Other teachers consistently expressed the opinion to the investigator that their descriptions of the teacher's behavior with pupils would be invalid. They believed that they were only guessing. Principals never expressed this opinion. Many other teachers stated that their descriptions would not agree with those descriptions by other judges and it was unnecessary to prove this true. With this belief the investigator tended to agree, but it was not substantiated by the results. To the contrary, the correlations

between the principals and the other teachers proved the highest on both dimensions. There are two possible reasons for this. Firstly, the principals might not have observed the teacher's classroom behavior any more than the other teachers. Very few principals make a practice or have the time to visit the classroom. Secondly, and this is supported by the correlations between students and other teachers, classroom leader behavior of the teacher may be inferred from general contact with the teacher. The type of behavior measured by the TLBDQ may be similar to the behavior that a teacher exhibits in his association with other people involved in the same type of work. This latter explanation has been discussed previously. It was believed that Consideration might be more easily inferred from general contact with the teacher than would Initiating Structure.

It was concluded that there was a considerable amount of disagreement between self-descriptions and those provided by the other judges. A similar disagreement was reported by Hemphill and Coons (2, p. 19). There was a significant correlation between students' descriptions of the teacher self-descriptions. Since there was no significant correlation between effectiveness ratings by other judges and self-ratings as reported in Chapter V, it might be concluded that there was less of the motivational factor affecting their descriptions

of themselves than there was affecting their effectiveness ratings of themselves. Teachers might be more objective when describing themselves than when rating their effectiveness.

REFERENCES

- (1) Halpin, A. W. Manual for the Leader Behavior Description Questionnaire. Columbus: Bureau of Business Research, The Ohio State University, 1957.
- (2) Hemphill, J. K., and A. E. Coons. "Development of the Leader Behavior Description Questionnaire," Leader Behavior: Its Description and Measurement, R. M. Stogdill and A. W. Coons, editors. Pp. 6-38. Research Monograph Number 88. Columbus, Ohio: Bureau of Business Research, The Ohio State University, 1957.
- (3) Siegel, S. Nonparametric Statistics for the Behavioral Sciences. Toronto: McGraw-Hill Book Co., 1956.

CHAPTER VII

THE RELATIONSHIP BETWEEN EFFECTIVENESS RATINGS AND DESCRIPTIONS OF LEADER BEHAVIOR

I. RESULTS AND ANALYSIS

High and Low Effectiveness Groups

The most important analysis of the data involved selecting the most effective and most ineffective teachers and comparing the behavior scores of these two groups. This was done to test the first hypothesis which stated that, Students and superiors (principals and superintendents) rate those teachers highest in teaching effectiveness who score above the median on both of the leader behavior dimensions, Consideration and Initiating Structure. Students and superiors rate those teachers lowest in teaching effectiveness who score below the median on both of the leader behavior dimensions.

To select for each judge the upper and lower groups in effectiveness, the cut off point was determined by examining the effectiveness ratings to try and choose approximately the same number of teachers for both groups. However, for some judges there was a large number of ties and this meant that some effectiveness groups were very unequal.

The most effective teachers were then categorized according to their scores on the behavior dimensions. If their scores on both dimensions were above the median, they were placed in one category, and if their scores on both dimensions were below the median, they were placed in a second category. The same categorization was carried out for the group of ineffective teachers. The results are summarized for each judge in Table X. The number of teachers who did not fall into either of these categories (above the median on both dimensions or below the median on both dimensions) may be determined from the difference between the total in the group and the total in the two categories. For example, in the case of judgments by students, 14 teachers were in the most effective group and 13 of these scored either above the median on both dimensions or below the median on both dimensions. Thus, only 1 of these most effective teachers was above the median on one dimension while below the median on the other dimension.

Fisher's Method for Exact Probability (1, p. 96) was applied to the data and Table X lists the probabilities that such distributions would happen by chance.

The results indicated that there was a significant relationship between extremes in effectiveness ratings and behavior scores, except for the judgments by teachers

TABLE X

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
 SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
 AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS

Effectiveness ratings and descriptions provided by		Above median on both dimensions		Below median on both dimensions		Probability (one-tailed)
Students-	Upper 14(24%) on effectiveness	11		2		.00003
	Lower 14(24%) on effectiveness		0	11		
Other teachers-	Upper 14(21%) on effectiveness	8		1		.0002
	Lower 11(19%) on effectiveness	0		9		
Principals-	Upper 30(51%) on effectiveness	19		2		.0000001
	Lower 13(22%) on effectiveness	0		13		
Superintendents-	Upper 23(39%) on effectiveness	13		1		.0000025
	Lower 13(22%) on effectiveness	0		11		
Teachers-	Upper 9 (15%) on effectiveness	5		1		not significant
	Lower 12(20%) on effectiveness	2		5		

NOTE: The significance level was $p = .01$.

themselves. The probabilities that such distributions would happen by chance were extremely small. Observation of the significant relationships showed that there was not one case of a teacher being in the lowest effectiveness group and scoring above the median on both dimensions. Most teachers who were judged as being ineffective exhibited behavior which was described as below the median on both dimensions. Nearly all the teachers who were judged as being effective exhibited behavior which was described as above the median on both dimensions of leadership. The first hypothesis was substantiated at this stage of the analysis.

A further analysis of the same data was carried out by utilizing behavior scores and effectiveness ratings provided by different judges. The purpose of this procedure was to eliminate the possibility of contamination as a factor causing the relationship and to provide a further test of the first hypothesis. Contamination meant that a judge would tend to answer both instruments in the same manner. That is the relationship between effectiveness groups and behavior descriptions might have been caused by this contamination factor rather than being an indication of a true relationship. When separate judges for each instrument were used, contamination was eliminated. Table XI summarized the results when the behavior descriptions were provided by students and effectiveness rating by judges other than students. Table XII summarized the results when the behavior descriptions were provided by

other teachers, Table XIII when behavior descriptions were provided by principals, Table XIV when behavior descriptions were provided by superintendents, and Table XV when behavior descriptions were provided by the teachers themselves.

The result of the analysis of this rearrangement of the data was that behavior descriptions provided by other teachers (Table XII) did not appear to be significantly related to the most effective and ineffective ratings as judged by the other sets of raters. The only exception to this was for effectiveness ratings provided by principals.

For students and superintendents the distributions were significant except for the ones involving teachers themselves, but the probabilities that such distributions would happen by chance were increased over the probabilities prior to the rearrangement.

For principals the distributions were significant when superintendents and other teachers provided the ratings of effectiveness. The probabilities that the distributions would happen by chance were increased from the probabilities indicated in Table X. The principals' descriptions resulted in the strongest probabilities that the distributions would not happen by chance.

Table XVI summarized the results for the half-groups of students. One half-group of students provided the ratings

TABLE XI

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
 SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
 AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS
AS DESCRIBED BY STUDENTS

Effectiveness ratings provided by	Above median on <u>both</u> dimensions	Below median on <u>both</u> dimensions	Probability (one-tailed)
Other teachers-	Upper 14(24%) on effectiveness	8	2
-	Lower 11(19%) on effectiveness	0	.001
Principals-	Upper 30(51%) on effectiveness	11	9
-	Lower 13(22%) on effectiveness	0	.008
Superintendents-	Upper 23(39%) on effectiveness	10	7
-	Lower 13(22%) on effectiveness	0	.006
Teachers-	Upper 9 (15%) on effectiveness	2	5
-	Lower 12(20%) on effectiveness	3	5
			not significant

NOTE: The significance level was $p = .01$.

TABLE XII

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
 SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
 AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS
AS DESCRIBED BY OTHER TEACHERS

Effectiveness ratings provided by		Above median on both dimensions	Below median on both dimensions	Probability (one-tailed)
Students-	Upper 14(24%) on effectiveness	7	2	
	Lower 14(24%) on effectiveness	2	7	not significant
Principals-	Upper 30(51%) on effectiveness	11	6	
	Lower 13(22%) on effectiveness	1	10	.005
Superintendents-	Upper 23(39%) on effectiveness	8	6	
	Lower 13(22%) on effectiveness	1	9	not significant
Teachers-	Upper 9 (15%) on effectiveness	2	3	
	Lower 12(20%) on effectiveness	3	3	not significant

NOTE: The significance level was $p = .01$.

TABLE XIII

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
 SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
 AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS
 AS DESCRIBED BY PRINCIPALS

Effectiveness ratings provided by		Above median on both dimensions	Below median on Probability both dimensions (one-tailed)
Students-	Upper 14(24%) on effectiveness	8	4
	Lower 14(24%) on effectiveness	1	7
Other teachers-	Upper 14(24%) on effectiveness	9	1
	Lower 11(19%) on effectiveness	0	9
Superintendents-	Upper 23(39%) on effectiveness	13	4
	Lower 13(22%) on effectiveness	0	11
Teachers-	Upper 9 (15%) on effectiveness	3	4
	Lower 12(20%) on effectiveness	4	5

NOTE: The significance level was $p = .01$.

TABLE XIV

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS
AS DESCRIBED BY SUPERINTENDENTS

Effectiveness ratings provided by		Above median on <u>both dimensions</u>		Below median on <u>both dimensions</u>	Probability (one-tailed)
Students-	Upper 14(24%) on effectiveness	9		4	.01
	Lower 14(24%) on effectiveness	1		8	
Other teachers-	Upper 14(24%) on effectiveness	8		3	.009
	Lower 11(19%) on effectiveness	1		8	
Principals-	Upper 30(51%) on effectiveness	13		5	.004
	Lower 13(22%) on effectiveness	2		10	
Teachers-	Upper 9 (15%) on effectiveness	2		5	not significant
	Lower 12(20%) on effectiveness	3		8	

NOTE: The significance level was $p = .01$.

THE STATE

THE STATE OF NEW YORK
 OFFICE OF THE ATTORNEY GENERAL
 ALBANY, N. Y.

IN SENATE,
 January 1, 1901.

REPORT OF THE
 ATTORNEY GENERAL

No.	Name	Age	Sex	Color	Religion	Education	Occupation	Married	Children	Notes
1	John A. Smith	45	M	W	Episcopal	High School	Teacher	Yes	2	
2	John B. Smith	35	M	W	Episcopal	High School	Teacher	Yes	2	
3	John C. Smith	25	M	W	Episcopal	High School	Teacher	Yes	2	
4	John D. Smith	15	M	W	Episcopal	High School	Teacher	Yes	2	
5	John E. Smith	10	M	W	Episcopal	High School	Teacher	Yes	2	
6	John F. Smith	5	M	W	Episcopal	High School	Teacher	Yes	2	
7	John G. Smith	4	M	W	Episcopal	High School	Teacher	Yes	2	
8	John H. Smith	3	M	W	Episcopal	High School	Teacher	Yes	2	
9	John I. Smith	2	M	W	Episcopal	High School	Teacher	Yes	2	
10	John J. Smith	1	M	W	Episcopal	High School	Teacher	Yes	2	

TABLE XV

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
 SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
 AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS
AS DESCRIBED BY TEACHERS THEMSELVES

Effectiveness ratings provided by		Above median on both dimensions	Below median on both dimensions	Probability (one-tailed)
Students-	Upper 14(24%) on effectiveness	9	3	not significant
	Lower 14(24%) on effectiveness	4	6	
Other teachers-	Upper 14(24%) on effectiveness	6	3	not significant
	Lower 11(19%) on effectiveness	3	7	
Principals-	Upper 30(51%) on effectiveness	13	8	not significant
	Lower 13(22%) on effectiveness	3	8	
Superintendent-	Upper 23(39%) on effectiveness	12	5	not significant
	Lower 13(22%) on effectiveness	3	7	

NOTE: The significance level was $p = .01$.

from which the most effective and least effective teachers were selected. The other half-group of students provided the behavior descriptions. Although both behavior descriptions and effectiveness ratings were provided by students, they were not the same group of students. This method of analysing the data also eliminated contamination. The probability was significant, but increased, that such a distribution would happen by chance.

The results of the analysis of the rearrangement of the data confirmed the conclusion that the first hypothesis was substantiated.

Rank-order Correlations

Rank-order correlations were calculated between effectiveness ratings and behavior descriptions using the formula for tied ranks. Table XVII summarized the correlations when the effectiveness ratings and the scores on each behavior dimension were provided by the same set of judges.

All the correlations, except that between the teachers' self-ratings and self-descriptions of Initiating Structure, were significant. There were significant relationships between teaching effectiveness and descriptions of both leader behavior dimensions for all judges except the teachers themselves.

TABLE XVI

NUMBER OF TEACHERS IN HIGH AND LOW EFFECTIVENESS GROUPS
 SCORING ABOVE THE MEDIAN ON BOTH LEADER BEHAVIOR DIMENSIONS,
 AND SCORING BELOW THE MEDIAN ON BOTH BEHAVIOR DIMENSIONS
 AS DESCRIBED BY ONE HALF-GROUP OF STUDENTS

Effectiveness ratings provided by other half-group of students	Descriptions by one half-group of students		Probability (one-tailed)
	Above median on <u>both</u> dimensions	Below median on <u>both</u> dimensions	
Upper 14(24%) on effectiveness	9	2	.001
Lower 16(27%) on effectiveness	2	12	

NOTE: The significance level was $p = .01$.

TABLE XVII

CORRELATIONS BETWEEN EFFECTIVENESS RATINGS AND
BEHAVIOR DESCRIPTIONS BY THE SAME JUDGE

Judge	Correlations between:		Significance of difference
	Initiating Structure and rating	Consideration and rating	
Students	.57**	.70**	not significant
Other teachers	.64**	.55**	not significant
Principals	.75**	.70**	not significant
Superintendents	.67**	.65**	not significant
Teachers	.18	.28*	not significant

NOTE: * significant at the .05 level.

** significant at the .01 level.

APPENDIX

THE LIST OF THE NAMES OF THE VESSELS WHICH WERE USED IN THE RESEARCHES OF THE ROYAL NAVY IN THE YEAR 1881

NAME OF THE VESSEL	CLASS OF VESSEL		NAME OF THE COMMANDER	NAME OF THE SURVEYOR	NAME OF THE ASSISTANT SURVEYOR
	TYPE	NUMBER			
ALBATROSS	Steamer	1	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	2	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	3	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	4	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	5	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	6	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	7	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	8	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	9	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns
ALBATROSS	Steamer	10	Mr. J. H. Mearns	Mr. J. H. Mearns	Mr. J. H. Mearns

THE LIST OF THE NAMES OF THE VESSELS WHICH WERE USED IN THE RESEARCHES OF THE ROYAL NAVY IN THE YEAR 1881

Between the effectiveness ratings and the Initiating Structure scores the greatest amount of significant relationship was provided by the principals, and the least amount of significant relationship was provided by the students. Between the effectiveness ratings and the Consideration scores the greatest amount of significant relationship was again provided by the principals, but it was equal to the relationship shown by the students. The least amount of significant relationship was provided by the teachers themselves.

The second hypothesis stated that, Students tend to relate effectiveness of teaching with teachers' scores on both Consideration and Initiating Structure, but Consideration is more closely related.

It was concluded that students did relate effectiveness to both behavior dimensions as shown by the significant correlations in Table XVII. To check whether Consideration was more closely related to effectiveness ratings than was Initiating Structure, the significance of the difference of the correlations was tested by making the z' transformation. The same criticism as was made previously, when a parametric test was used, applied here. Although the difference appeared quite large and in the hypothesized direction, it was not significant at the .05 level.

A further check of this hypothesis was provided through an analysis of the data pertaining to the half-groups of students. One half-group was used to provide effectiveness ratings and the other half-group was used to provide behavior descriptions. The correlations were .42 between Initiating Structure and the effectiveness ratings and .47 between Consideration and the effectiveness ratings. Both correlations were significant at the .01 level, but the difference between them was not significant. This agreed with the conclusion in the above paragraph-- for students, effectiveness was related to both dimensions, but not more closely related to Consideration.

A third check of this hypothesis was provided by correlations between effectiveness ratings by students and behavior descriptions by judges other than students. When behavior descriptions were provided by principals and the effectiveness ratings by the students, only the correlation involving Consideration was significantly related as may be noted from Table XVIII. There was, though, a significant difference at the .05 level between the correlations. Only one dimension was significantly related to effectiveness ratings, and therefore, Consideration was more closely related.

TABLE XVIII

CORRELATIONS BETWEEN EFFECTIVENESS RATINGS AND
BEHAVIOR DESCRIPTIONS BY SEPARATE JUDGES

Behavior descriptions provided by:		Ratings provided by:		
		Students	Principals	Superintendents
Students-	I. S.	-	.40**	.33**
	- C.		.45**	.34**
Principals-	I. S.	.25	-	.49**
	- C.	.55**		.48**
Superintendents-	I. S.	.30*	.53**	-
	- C.	.57**	.49**	.

NOTE: * significant at the .05 level.

** significant at the .01 level.

I. S.- Initiating Structure.

C. - Consideration.

When behavior descriptions were provided by superintendents, both dimensions were significantly related to effectiveness ratings by students. The difference between them was significant at the .05 level in the hypothesized direction, indicating that Consideration was more closely related to effectiveness ratings by students.

The third hypothesis stated that, Superiors (principals and superintendents) tend to relate effectiveness of teaching with teachers' scores on both Consideration and Initiating Structure, but Initiating Structure is more closely related.

As observed in Table XVII, when the principals provided the effectiveness ratings and the descriptions, both dimensions were significantly related to effectiveness ratings, but there was no significant difference. This same result is observed in Table XVIII when the superintendents and students provided the behavior descriptions and the principals the effectiveness ratings. Thus, for principals, both dimensions of teacher leader behavior were related to effectiveness ratings, but Initiating Structure was not more closely related.

As observed in Table XVII, when the superintendents provided the effectiveness ratings and the descriptions, both dimensions were significantly related to effectiveness ratings, but there was no significant difference between them. This same result is observed in Table XVIII when the students and

principals provided the behavior descriptions and the superintendents the effectiveness ratings. Thus, superintendents related teaching effectiveness to both dimensions of teacher leader behavior, but Initiating Structure was not more closely related.

Other teachers significantly related teaching effectiveness to behavior on both dimensions and there was no significant difference in the correlations. Teachers themselves related teaching effectiveness only to Consideration. These results may be observed in Table XVII.

It should be noted that all the significant correlations discussed in this section indicated partial relationship and not perfect relationship.

The correlations might be a result of the contamination factor explained previously. The relationships might be due to judges responding to both instruments in a like manner. This factor would be eliminated when separate judges were used to describe and rate. In Table XVIII the results indicated that there was still a relationship when contamination had been eliminated, although the relationships were lower. Contamination was also eliminated when the students were split into two half-groups. The correlation was .42 between Initiating Structure and the effectiveness ratings for the half-groups of students, and .57 between Initiating Structure and

effectiveness ratings for the whole group of students. The difference between them was not significant. Between Consideration and the effectiveness ratings the correlations were .47 for the half-groups of students and .70 for the whole group of students. The difference between them was significant at the .05 level. Contamination apparently was a factor with Consideration, not with Initiating Structure, but even with it eliminated the correlations were significant.

II. DISCUSSION

This chapter reported and analysed the data used to test the three hypotheses. The first hypothesis stated that those teachers who were most effective would be above the median on both dimensions and those teachers who were least effective would be below the median on both dimensions. This was supported. All judges, except the teachers themselves, confirmed this result. When separate judges were utilized for behavior descriptions and effectiveness ratings, the result was still supported for behavior descriptions by students, principals, and superintendents although the probabilities that such distributions would happen by chance were increased.

The method of analysis that was used allowed a tentative conclusion that the extent of teacher leader behavior

was related to effective or ineffective teaching. When leader behavior was thought to be present to a greater than average extent, the teachers were effective; when leader behavior was thought to be present to a less than average extent, the teachers were ineffective.

The second hypothesis stated that students would relate effectiveness ratings to behavior scores on both dimensions, but Consideration would be more closely related. For behavior descriptions and effectiveness ratings by the students, both dimensions were related and there was a difference (but not significant) in the hypothesized direction. For the half-groups of students, both dimensions were related to effectiveness ratings and there was no significant difference between them. When effectiveness ratings provided by the students were correlated with descriptions provided by the superintendents, both dimensions were significantly related and Consideration was more closely related. In the case of behavior descriptions provided by the principals, only that between Consideration and the effectiveness ratings was significant. The general conclusion was that students related effectiveness to scores on both dimensions and there was a slight tendency to relate their ratings of effectiveness more closely to Consideration.

The final hypothesis stated that both behavior

dimensions would be related to effectiveness ratings by principals and superintendents, but that Initiating Structure would be more closely related. For both principals and superintendents there was a significant relationship between their effectiveness ratings and behavior descriptions provided by themselves, students and the other administrator, but Initiating Structure was not more closely related. This latter conclusion was unexpected as it was believed that administrators preferred the teacher who got things done. In this study administrators considered the teacher's human relations behavior in the classroom as important a part of effective teaching as his Initiating Structure behavior.

The use of separate judges for behavior descriptions and effectiveness ratings eliminated the contamination factor. When contamination was eliminated, the relationships were decreased, but remained significant. For the students' half-groups, contamination significantly reduced the correlation between Consideration and effectiveness ratings, but not the correlation between Initiating Structure and effectiveness ratings. In the case of students the tendency for judges to respond to both instruments in the same manner appeared to be a more important factor between Consideration and effectiveness than it did between Initiating Structure and effectiveness. It might be that general "likes" and "dislikes", which undoubtedly affect

a judge's response on the global rating, have a greater effect on his perception of a teacher's human relations behavior than the effect they have on his perception of a teacher's Initiating Structure behavior.

REFERENCES

- (1) Siegel, S. Nonparametric Statistics for the Behavioral Sciences. Toronto: McGraw-Hill Book Co., 1956.

CHAPTER VIII

CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

An attempt has been made to study elements of behavior common to teachers. Empirical research in other fields has suggested that dimensions of leader behavior might be applicable to teacher-student relationships. The relationships between these dimensions of Consideration and Initiating Structure, and ratings of teaching effectiveness have been investigated for fifty-nine teachers in twelve schools in the Province of Alberta. Five viewpoints or frames of reference have been utilized to provide global effectiveness ratings and descriptions of teacher leader behavior; they were students, other teachers on the same staff, principals, superintendents, and the teachers themselves.

The only evidence provided that the conclusions could be generalized beyond the persons in the study was the wide variation of experience, training, grades and subjects taught, background, and schools of the fifty-nine teachers. No urban schools or elementary teachers were used in the core of the study.

All the participating teachers volunteered. Thus, many schools were not included because one or more of the members of the staff did not wish to take part. The study was conducted

under the condition that the results were not to be used for any administrative purposes in the schools in which the data were collected. Complete anonymity of the participants was maintained.

The Teacher Leader Behavior Description Questionnaire was used to provide descriptions of teacher leader behavior. A global rating scale provided the ratings of teaching effectiveness.

I. CONCLUSIONS

1. The most important conclusion was that part of teaching effectiveness appeared to be related to the extent of a teacher's Consideration and Initiating Structure behavior. The most effective teachers exhibited behavior that was described as being above the median on both dimensions while the least effective teachers exhibited behavior that was described as being below the median on both dimensions.

2. Students appeared to relate teaching effectiveness to the teacher's leader behavior and there was some indication that they believed human relations behavior (Consideration) was more closely related to effectiveness than was goal accomplishment behavior (Initiating Structure).

3. Principals and superintendents related teaching effectiveness to the teacher's leader behavior. Neither of these administrators considered that the teacher's behavior in trying to attain goals was any more closely related to effectiveness than was the teacher's behavior in human relations.

4. The teachers themselves as judges were in a category distinct from all the other judges. They showed very little, if any, agreement with other judges on ratings of effectiveness or on behavior descriptions. They did not show any relationship between most effective and most ineffective ratings and behavior descriptions. It was suggested that spurious factors affected their judgments to a greater extent than it affected judgments by others.

5. Since the amount of agreement on both behavior dimensions was not greatly different from the amount of agreement on effectiveness ratings, it was concluded that it was just as difficult to agree on certain aspects of what the teacher does in the classroom as it was to agree on how well he teaches in general.

6. The principals and superintendents probably have had the most experience in evaluating teachers whether it be formal or informal evaluation. They probably have had the

most supervisory experience with teachers where attempts have been made to understand and help individual teachers to be more effective. Thus, they appeared to be in a more similar position than any other pair of judges with regard to understanding teaching effectiveness and they showed the most agreement on the ratings of effectiveness.

The superintendents and students showed the least agreement on the effectiveness ratings (excluding self-ratings). They were likely in the most dissimilar position for developing a frame of reference to judge the effectiveness of a teacher.

7. The principals and other teachers showed the greatest amount of agreement on both behavior dimensions. They were probably in the most similar position when it came to describing the teacher's everyday behavior in the classroom.

8. It was concluded that an administrator might be able to infer a teacher's human relations behavior in the classroom from personal contact with him outside of the classroom. It was not considered as possible that an administrator would be able to infer a teacher's classroom Initiating Structure behavior from personal contact with the teacher outside of the classroom.

9. The students were the only judges who were constantly in contact with the teacher and they were the only judges who showed some significant agreement on descriptions of behavior with all other judges including the teachers themselves.

10. Conclusions regarding the use of the TLBDQ as an instrument to describe teacher leader behavior were indefinite. The reliability of the Initiating Structure dimension was not considered satisfactorily high, and the independence of the two dimensions was lower than desirable. One must conclude that the present TLBDQ should not be used by any one set of judges and that all five viewpoints used in this study should be considered.

II. IMPLICATIONS

The results of this research may be useful to those persons concerned with the improvement of the most critical aspect of the educational system, the instructional personnel.

1. Administrators will find it valuable to know that there was some relationship between a teacher's effectiveness in the classroom and his leader behavior. Of special significance was the conclusion that the most effective teachers were above the median in both Consideration and Initiating Structure

and that ineffective teachers were below the median on both dimensions. The emphasis on one dimension to the neglect of the other might result in an unsatisfactory solution to a particular problem. An understanding and study of leader behavior will afford a means whereby new and ineffective teachers might be helped without the use of any instrument. The study of leader behavior might prove valuable for in-service education programs.

2. There was evidence in this research that teacher leader behavior, as studied through the TLBDQ, did not account entirely for teacher effectiveness or ineffectiveness. Firstly, behavior of the teacher which is a product of his technical knowledge such as skill in subject matter might be a factor. This would vary from grade to grade and subject to subject. Secondly, effective teacher behavior might be a function of the situational variables, such as adaptation to a class of slow learners.

If the above is true, then administrators must realize that leader behavior, behavior applicable to the situation, and behavior reflecting a knowledge of subject matter must all be considered in a complete analysis of effectiveness. Reliance on one or two of the factors to the neglect of the other will result in an incomplete analysis of a problem.

3. Administrators need to broaden their frame of reference for perceiving teachers. They need especially to take into account the student's relationships with the teacher.

Superintendents could use the principals' descriptions of what the teachers do, to supplement their own, since their descriptions were at variance with those of the principals.

Superintendents and principals need to obtain a more complete picture of how the teacher provides and maintains relationships to get things done and to accomplish goals.

With regard to obtaining a broad picture of the teacher's leader behavior in the classroom, administrators can provide no more agreement among themselves, or with students, than their agreement with those other teachers who have had no opportunity to observe the teacher's classroom behavior. This emphasizes a fact, about which most administrators are well aware, that their visitations just scratch the surface of obtaining a complete picture of the teacher's performance.

4. Administrators need to communicate with teachers the perception that they have of the teacher's effectiveness and behavior. Also, they need to obtain the teacher's

perception of himself.

5. If leader behavior is stable, the results may have implications for teacher selection prior to enrolling in a teachers' college. If leader behavior is not stable and changes can be produced, then the results may have implications for courses on a teacher education program.

6. There are also implications from this research for the selection of administrators. Other things being equal, the most effective teacher could be educated to be the most effective administrator as he would have exhibited behaviors necessary for successful leadership.

III. RECOMMENDATIONS

1. Study is recommended to enlarge the understanding of leader behavior and specifically understanding of the two dimensions. More complete knowledge is needed of what other behavior items, aside from specific items on the TLBDQ, can be classified as Consideration or Initiating Structure. What behavior definitely cannot be classified under these two dimensions? How does an individual develop these behavior patterns and is it possible to make any significant change in them?

2. The conclusion that teacher leader behavior is to some extent related to teaching effectiveness needs to be confirmed in other situations especially in urban centres and with elementary teachers.

3. To make the TLBDQ a more effective instrument a further check should be conducted of the reliability of the dimensions. An item analysis should be carried out to eliminate items which do not show a relationship to their respective dimensions, and to reduce the dependence between the two dimensions.

An analysis of the use of junior high school students compared to senior high school students as respondents might be worthwhile.

4. A study is recommended of the relation between Consideration and Initiating Structure and the productivity of the pupils, and also the relation between Consideration and Initiating Structure and the teacher predictor variables such as education, knowledge, and intelligence.

5. In the field of teacher effectiveness some study is recommended of the relationship of the situational variables to teacher behavior and the relationship of the situational variables to pupil productivity.

B I B L I O G R A P H Y

BIBLIOGRAPHY

A. BOOKS

- Beecher, D. W. The Evaluation of Teaching. Syracuse: Syracuse University Press, 1949.
- Burton, W. H., and L. J. Brueckner. Supervision a Social Process. Third Edition. New York: Appleton-Century-Crofts Inc., 1955.
- Monroe, W. S. (ed.). Encyclopedia of Educational Research. New York: The MacMillan Co., 1952.
- Morris, B. H. Personal Traits and Success in Teaching. Teachers College Contributions to Education, No. 342. New York: Bureau of Publications, Teachers College, Columbia University, 1929.
- Payne, E. G. (ed.). Readings in Educational Sociology, Vol. II. New York: Prentice-Hall Inc., 1940.
- Reavis, W. C., and D. H. Cooper. Evaluation of Teacher Merit in City School Systems. Chicago: University of Chicago Press, 1945.
- Shane, H. G., and W. A. Yauch. Creative School Administration. New York: Henry Holt and Co., 1954.
- Sherif, M., and H. Cantril. The Psychology of Ego-Involvements. New York: John Wiley and Sons, 1947.
- Siegel, S. Nonparametric Statistics for the Behavioral Sciences. Toronto: McGraw-Hill Book Co., 1956.
- Young, L. P. The Administration of Merit-Type Teacher Salary Schedules. New York: Teachers College, Columbia, Columbia University, 1933.

B. PERIODICALS, MONOGRAPHS AND UNPUBLISHED WORKS

Ackerman, W. I. "Teacher Competence and Pupil Change," Harvard Educational Review, 24:273-289, 1954.

Andrews, J. H. M. "Administrative Significance of Teacher Differences," Unpublished Doctoral dissertation, The University of Chicago, 1957.

Biber, B. "Problems of Values and Measures in Evaluation of Teaching," Educational Leadership, 15:213-217, January, 1958.

Gage, N. L., and J. S. Orleans. "Guiding Principles in the Study of Teacher Effectiveness," Journal of Teacher Education, 3:294-298, December, 1952.

Gray, R. "Teacher Evaluation," Canadian Education, 10:171, June, 1955.

Halpin, A. W. "The Behavior of Leaders," Educational Leadership, 14:172-176, December, 1956.

_____. Manual for the Leader Behavior Description Questionnaire. Columbus: Bureau of Business Research, The Ohio State University, 1957.

_____. Studies in Aircrew Composition. III. The Combat Leader Behavior of B-29 Aircraft Commanders. Bolling Air Force Base, Washington, D. C.: Human Factors Operations Research Laboratories, 1953, (HFORL Memorandum No. TN-54-7).

Hemphill, J.K. "Leader Behavior Associated with the Administrative Reputations of College Departments," Leader Behavior: Its Description and Measurement, (R. M. and A. E. Coons, editors). Pp. 74-102. Research Monograph Number 88. Columbus, Ohio: Bureau of Business Research, The Ohio State University, 1957.

Hemphill, J. K., and A. E. Coons. "Development of the Leader Behavior Description Questionnaire," Leader Behavior: Its Description and Measurement, (R. M. Stogdill and A. E. Coons, editors). Pp. 6-38. Research Monograph Number 88. Columbus, Ohio: Bureau of Business Research, The Ohio State University, 1957.

- Hughes, M. M. "Whither Evaluation?" Educational Leadership, 15:208-212, January, 1958.
- Kinney, L. B. "What Is a Good Teacher?" School of Education Bulletin. Utah State Agricultural College, Utah, 1954.
- Levin, H. "A New Perspective on Teacher Competence Research," Harvard Educational Review, 24:98-105, 1954.
- Melby, E. O. "Role of Evaluation in Improving Teaching," Educational Leadership, 15:218-220, January, 1958.
- Mitzel, H. E. A Behavioral Approach to the Assessment of Teacher Effectiveness, Monograph. New York: Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, February, 1957. (mimeographed).
- Orleans, J. S., et al. "Some Preliminary Thoughts on the Criteria of Teacher Effectiveness," Journal of Educational Research, 45:641-648, May, 1952.
- Rose, G. "Toward the Evaluation of Teaching," Educational Leadership, 15:231-238, January, 1958.
- Rugg, H. "Is the Rating of Human Character Practicable?" Journal of Educational Psychology, 12:425-438, November, 1921; 12:485-501, December, 1921; 13:30-42, January, 1922.
- Ryans, D. G. (chairman), et al. "Report of the Committee on the Criteria of Teacher Effectiveness," Review of Educational Research, 22:238-263, June, 1952.
- Scott, E. L. Leadership and Perceptions of Organization. Research Monograph Number 82. Columbus: The Bureau of Business Research, The Ohio State University, 1956.
- Taft, R. "The Ability to Judge People," Psychological Bulletin, 52:1-23, January, 1955.
- Thayer, V. T. "Teacher Rating in the Secondary School," Educational Administration and Supervision, 12:361-378, September, 1926.
- Wandt, R., and L. M. Ostreicher. Variability in Observed Classroom Behaviors of Junior High School Teachers and Classes. Research Monograph Number 16. New York: Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, 1953.

A P P E N D I C E S

APPENDIX A

QUESTIONNAIRES AND DIRECTIONS

TEACHER LEADER BEHAVIOR DESCRIPTION QUESTIONNAIRE

On the following pages is a list of items that may be used to describe the behavior of teachers in their work with students. Each item describes a specific kind of behavior, but does not ask you to judge whether the behavior is desirable or undesirable. This is not a test of ability. It simply asks you to describe, as accurately as you can, the behavior of a certain teacher.

DIRECTIONS:

- a. READ each item carefully. Read "she" for "he" if the teacher being described is female.
- b. THINK about how frequently the teacher engages in the behavior described by the item.
- c. DECIDE whether he always, often, occasionally, seldom, or never acts as described by the item.
- d. DRAW A CIRCLE around one of the five letters following the item to show the answer you have selected using the following:
A - Always, B - Often, C - Occasionally, D - Seldom, E - Never.

1. He helps students with their personal problems. ABCDE
2. He makes his attitudes clear to the class. ABCDE
3. He does little things to make it pleasant to be a member of the class. ABCDE
4. He tries out his new ideas with the class. ABCDE
5. He is easy to understand. ABCDE
6. He rules with an iron hand. ABCDE
7. He finds time to listen to students. ABCDE
8. He criticizes poor work. ABCDE
9. He speaks in a manner not to be questioned. ABCDE
10. He talks with students outside of class. ABCDE
11. He looks out for the personal welfare of individual class members. ABCDE
12. He assigns class members to particular tasks. ABCDE
13. He schedules the work to be done. ABCDE
14. He maintains definite standards of performance. ABCDE
15. He gives students reasons or explanations for his actions. ABCDE
16. He acts without consulting the class. ABCDE
17. He praises students. ABCDE
18. He emphasizes the meeting of deadlines. ABCDE
19. He treats students fairly. ABCDE
20. He encourages the use of uniform procedures. ABCDE
21. He is willing to make changes. ABCDE

- | | |
|---|-------|
| 22. He makes sure that his job as teacher is understood by all class members. | ABCDE |
| 23. He is friendly and approachable. | ABCDE |
| 24. He asks that students follow standard rules and regulations. | ABCDE |
| 25. He makes students feel at ease when talking with him. | ABCDE |
| 26. He lets class members know what is expected of them. | ABCDE |
| 27. He puts suggestions made by the class into operation. | ABCDE |
| 28. He sees to it that class members are working up to capacity. | ABCDE |
| 29. He gets students to help in the planning of class activities. | ABCDE |
| 30. He sees to it that the work of the class is organized. | ABCDE |

Please make sure that you have answered every question.

NOTE: Item number 16 is scored negatively.

The Consideration items are numbers: 1, 3, 5, 7, 10, 11, 15, 16, 17, 19, 21, 23, 25, 27 and 29.

The Initiating Structure items are numbers: 2, 4, 6, 8, 9, 12, 13, 14, 18, 20, 22, 24, 26, 28 and 30.

GLOBAL RATING OF TEACHER EFFECTIVENESS

Please check the statement which in your judgment best indicates how well this teacher does in teaching his students.

PLEASE CHECK IN ONE SQUARE.

1. Outstanding for this school.
2. High for this school.
3. Slightly better than average for this school.
4. Slightly below average for this school.
5. Low for this school.
6. Very low for this school.

DIRECTIONS FOR ADMINISTERING INSTRUMENTS TO STUDENTS

1. This is a research project which is studying teaching. You will be asked to answer a questionnaire which will be distributed later.
2. This study will not mention names of school, teachers or students.
3. There was no particular reason why this school or teacher or you as students were chosen. It was all done by chance.
4. The study is not concerned with individuals. The results will be about a number of teachers in the province, not those in a particular school or division.

5. No one will see the answered questionnaires, but the investigator.
6. Do not put your name on the questionnaire.
7. Write the teacher's name on the slip attached at the top. This will be torn off after you have finished and a number has been placed on the questionnaire.
8. If you are in grades eight or nine place a "J" on the front of your paper.
9. Read the printed instructions on the questionnaire to the students.

DIRECTIONS FOR ADMINISTERING INSTRUMENTS TO
SUPERINTENDENTS, TEACHERS THEMSELVES,
PRINCIPALS AND OTHER TEACHERS

1. Write the name of the teacher being described on the slip which is attached. This will be removed after you have finished and a number has been placed on the questionnaire by the investigator.
2. Do not write your name on the questionnaire.
3. Answer the questions about the teacher to the best of your knowledge. You may have to infer what the teacher does in the classroom from your experience with him or her.
4. Answer the questions individually without referring to the other questions.

5. On the back page would you please rate his or her teaching of students using as a standard the teachers of this school. In your rating, fields of teacher activity outside of teaching students should not be included.

APPENDIX B

LETTER TO SUPERINTENDENTS

The following is a copy of the letter sent to provincially appointed superintendents of schools in school divisions and counties in the Province of Alberta. This letter was the means used to obtain the co-operation of the twelve schools used in this study.

Dear Sir:

This letter is to inquire whether you would be willing to assist in a research study which I am conducting as a graduate student in the Division of Educational Administration, University of Alberta.

The purpose of this study is to aid in the determination of teacher competence and to see if there are common patterns of effective performance of teachers.

The thesis is designed to study descriptions of teachers' classroom performance and to relate these descriptions to ratings of teacher effectiveness. It is hoped that it might provide a more thorough background for study of such problems as teacher education, in-service education and teacher selection by determining specific areas of performance which are related to effective teaching.

It is planned to carry out the study by using one

school in each of twelve Divisions. In each school five teachers teaching above grade six would be chosen at random at the start of the study. The intent is to have each teacher's classroom performance described by ten students, five other teachers, the principal, the superintendent, and the teacher himself, and also have his teaching rated by the same people. This would involve the use of two very simple instruments which could be completed in five minutes and no preparation prior to the administering of the instruments would be required.

To carry this out, you would be asked to do the following:

Obtain the co-operation of the principal, and the entire staff of one school in your Division. This school should have a minimum of ten teachers on the staff and at least eight should teach above grade six.

Nothing further would be necessary until the actual administering of the instruments by myself which I would hope could be done in the latter part of April or the first week of May.

Since some might be reluctant to take part in a study such as this the following should be noted:

I will administer the instruments myself and no one else will see or have access to them. The information will

be used only for research.

The reported study will involve no names either of individuals, schools or Divisions. There will be no attempt to compare schools or Divisions and no such breakdown will appear.

A list of the names of the five persons being described with corresponding numbers will be made up upon arrival at the school and be destroyed before leaving each school. The instruments will bear the teacher's number and his name will appear on an attached slip to serve only as identification, so that individuals will be describing the same teacher. The slips then will be destroyed; all persons taking part will remain anonymous.

This study has the approval of the Alberta Teachers' Association. I have been a member of the provincial executive of the Saskatchewan Teachers' Federation and am very concerned that all those taking part would do so willingly and that no school would be utilized in which a single person did not wish to do so.

I would appreciate an answer when you have been able to determine whether one of your schools would co-operate, and whether you would be willing. You could inform me as to the time which would be satisfactory.

I have enclosed a card to facilitate your reply. I

would be only too willing to provide additional information or to answer your questions.

Yours sincerely,

A. G. McBeath

B29783